

(19) World Intellectual Property Organization
International Bureau

01 OCT 2004

(43) International Publication Date
16 October 2003 (16.10.2003)

PCT

(10) International Publication Number
WO 03/085095 A2(51) International Patent Classification⁷: C12N

(21) International Application Number: PCT/US03/09921

(22) International Filing Date: 1 April 2003 (01.04.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

10/112,372	1 April 2002 (01.04.2002)	US
60/382,614	24 May 2002 (24.05.2002)	US
10/164,717	10 June 2002 (10.06.2002)	US
10/167,631	13 June 2002 (13.06.2002)	US
10/177,917	24 June 2002 (24.06.2002)	US
60/399,125	30 July 2002 (30.07.2002)	US

(71) Applicant (for all designated States except US): **ORIGENE TECHNOLOGIES, INC.** [US/US]; 6 Taft Court, Suite 100, Rockville, MD 20850 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **JAY, Gilbert** [US/US]; 5801 Nicholson Lane, North Bethesda, MD 20852 (US). **KOVACS, Karl, F.** [US/US]; 5 Gruenther Court, Rockville, MD 20851 (US). **LI, Xuan** [US/US]; 14808 Carona Drive, Silver Spring, MD 20905 (US). **FAN, Wufang** [US/US]; 10790 Roselle Street, San Diego, CA 92121 (US). **SHU, Youmin** [US/US]; 2508 Chilham Place, Potomac, MD 20854 (US). **YEE, Anthony** [US/US]; 3024 Bel Pre Road, No 101, Silver Spring, MD 20906 (US).(74) Agent: **LEBOVITZ, Richard, M.**; Origene Technologies, Inc., Suite 100, 6 Taft Court, Rockville, MD 20850 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),

Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

— as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

— of inventorship (Rule 4.17(iv)) for US only

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL EXPRESSED GENES

(57) Abstract: The present invention relates to all facets of novel polynucleotides, the polypeptides they encode, antibodies and specific binding partners thereto, and their applications to research, diagnosis, drug discovery, therapy, clinical medicine, forensic science and medicine, etc. The polynucleotides are useful in variety of ways, including, but not limited to, as molecular markers, as drug targets, and for detecting, diagnosing, staging, monitoring, prognosticating, preventing or treating, determining predisposition to, etc., diseases and conditions.



WO 03/085095 A2

20/pts.
-1-

DT09 Rec'd PCT/PTO 01 OCT 2004

NOVEL EXPRESSED GENES

This application claims the benefit of U.S. Serial No. 10/112,372, filed April 1, 2002, U.S. Serial No. 60/382, 614, filed May 24, 2002, U.S. Serial No. 10/164,717, filed June 10, 2002, U.S. Serial No. 10/167,631, filed June 13, 2002, U.S. Serial No. 10/177,917, filed June 24, 2002, and U.S. Serial No. 60/399,125, filed 30 July 2002, which are hereby incorporated by reference in their entirety.

DESCRIPTION OF THE DRAWINGS

Fig. 1 shows the expression of OTB0949 in human tissues. PCR was performed using SEQ NO 3 as the forward primer, and SEQ ID NO 4 as the reverse primer.

Fig. 2 shows the expression pattern of human OTB182, an integral membrane protein, in human tissues. To detect gene expression, PCR was carried out on aliquots of the normalized tissue samples using a forward (SEQ ID NO 13) and reverse (SEQ ID NO 14) gene-specific primers.

Fig. 3 shows the amino acid alignment of human OTB182 (SEQ ID NO 12) with mouse AK003645 (SEQ ID 15).

Fig. 4 (A-D) shows the amino acid alignments of a human transient receptor potential cation channel (TRPCC) with related channel family members. Human sequences are TRPM (SEQ ID NO 17), human AB046836 (SEQ ID NO 19), human XM_036123 (SEQ ID NO 18), and mouse XM_140575 (SEQ ID 20).

Fig. 5 shows the expression pattern of human TRPCC in human tissues. To detect gene expression, PCR was carried out on aliquots of the normalized tissue samples using a forward (SEQ ID NO 21) and reverse (SEQ ID NO 22) gene-specific primers.

Fig. 6 shows the amino acid sequence alignments between different forms of the human melanocortin-1 receptor. NM_002386 or MCR-1A (SEQ ID NO 30). MCR-1C (SEQ ID NO 26). MCR-1B (SEQ ID NO 31).

Fig. 7 shows a schematic of the exon sizes for the melanocortin-1 gene and the tubulin gene (exon 7).

Fig. 8 shows the expression pattern of OTB860 in human tissues. SEQ ID NOS 40 and 41 are the primer sequences.

Fig. 9 (A-C) shows the amino acid alignments of OTB860 (SEQ ID NO 39) and KIAA1678 (SEQ ID NO 42).

-2-

Fig. 10 (A and B) is the amino acid alignments of the different splice variants of human TARPP, Br137A (SEQ ID NO 46), B (SEQ ID NO 48), C (SEQ ID NO 50), D (SEQ ID NO 52; SEQ ID NO 14, NM_016300), and E (SEQ ID NO 44), and partial clone AL133109 (SEQ ID NO 55).

5 Fig. 11 is a schematic drawing showing the differences between the various forms of human TARPP.

Fig. 12 (A-C) shows amino acid alignments of the different splice variants of human TARPP (Br137A, B, C, D, and E) with mouse TARPP (NM_033264; SEQ ID NO 53).

The following procedure was used for the expression profile. A twenty-four tissue
10 panel was used (lanes from left to right): 1, adrenal gland; 2, bone marrow; 3, brain; 4, colon; 5, heart; 6, intestine; 7, pancreas; 8, liver; 9, lung; 10, lymph node; 11, lymphocytes; 12, mammary gland; 13, muscle; 14, ovary; 15, pancreas; 16, pituitary; 17, prostate; 18, skin; 19, spleen; 20, stomach; 21, testis; 22, thymus; 23, thyroid; 24, uterus. The lane at the far left of each panel contains molecular weight standards. The results were obtained according to
15 the following procedures:

Polyadenylated mRNA was isolated from tissue samples, and used as a template for first-strand cDNA synthesis. The resulting cDNA samples were normalized using beta-actin as a standard. For the normalization procedure, PCR was performed on aliquots of the first-strand cDNA using beta-actin specific primers. The PCR products were visualized on an
20 ethidium bromide stained agarose gel to estimate the quantity of beta-actin cDNA present in each sample. Based on these estimates, each sample was diluted with buffer until each contained the same quantity of beta-actin cDNA per unit volume.

To detect gene expression, PCR was carried out on aliquots of the normalized tissue samples using a forward and reverse gene-specific primers. The reaction products were
25 loaded on to an agarose (e.g., 1.5-2%) gel and separated electrophoretically.

DESCRIPTION OF THE INVENTION

The present invention relates to all facets of the novel genes described herein, polypeptides encoded by them, antibodies and specific binding partners thereto, and their
30 applications to research, diagnosis, drug discovery, therapy, clinical medicine, forensic science and medicine, etc. The polynucleotides and polypeptides are useful in variety of

-3-

ways, including, but not limited to, as molecular markers, as drug targets, and for detecting, diagnosing, staging, monitoring, prognosticating, preventing or treating, determining predisposition to, etc., diseases and conditions, associated with genes of the present invention. The identification of specific genes, and groups of genes, expressed in pathways physiologically relevant to particular tissues, permits the definition of functional and disease pathways, and the delineation of targets in these pathways which are useful in diagnostic, therapeutic, and clinical applications. The present invention also relates to methods of using the polynucleotides and related products (proteins, antibodies, etc.) in business and computer-related methods, e.g., advertising, displaying, offering, selling, etc., such products for sale, commercial use, licensing, etc.

OTB0949

OTB0949 is a polynucleotide (SEQ ID NO 1-2) which is expressed predominantly in brain tissue. Low levels of expression are observed in other tissues, e.g., adrenal gland, mammary gland, pituitary, stomach, and testes, but brain expression is at least 100-fold higher. See, e.g., Fig. 1. Because of its selectivity for brain, OTB0949 can be used as a molecular marker for brain tissue, e.g., in pathology and cytology, as well as a target, e.g., to ablate brain tissue, to deliver drugs to brain cells, etc. In the brain, OTB0949 is highly expressed in amygdala, hippocampus, thalamus, and retina. OTB0949 can also be a useful in diagnostics and therapeutics to treat neurological and visual disorders.

The brain is one of the most complicated and least understood organs in the mammalian body. Anatomically, it is composed of four different regions: (1) cerebral hemispheres, (2) diencephalon (thalamus, hypothalamus, and epithalamus), (3) brain stem (midbrain, pons, and medulla oblongata), and (4) cerebellum. These can be further subdivided. For instance, the cerebral hemispheres contain cerebral cortex and basal ganglia (caudate nucleus, putamen, globus pallidus, lentiform nucleus, corpus striatum, amygdala). The midbrain contains, e.g., cerebral peduncles, corpora quadrigemina, colliculi, substantia nigra, and the red nucleus. Other regions and subdivisions of interest include hypothalamus, pituitary, cranial nerves, pineal, gray matter, white matter, raphe nucleus, limbic system, etc. Various cell types are found in the brain, including, supporting cells, such as neuroglia, glia, astrocytes, microglia, ependymal cells, oligodendrocytes, and Schwann cells, neurons, such

as multipolar, bipolar, unipolar, Purkinje, and pyramidal cells.

The gene is located at chromosomal position 12q24.2. Several neurological diseases have been mapped to this region, e.g., spinocerebellar ataxia 2 (associated with a mutation in the ataxin-2 gene), spinal muscular atrophy (OMIM 158590), and amyotrophy (OMIM 181405). Disruption of OTB0949 (e.g., in the corresponding gene a transgenic animal) can result in a brain disorder or susceptibility thereto, including those mentioned above.

OTB0949 is coded for in a single exon, and comprises a short coding sequence of about 135 amino acids (SEQ ID NO 2). It contains a stretch of hydrophobic amino acids from about positions 76-100 (SEQ ID NO 2). Examples of promoters include, e.g., SEQ ID NOS 5-10 located in contig NT_009775 at about 557-607, 1263-1313, 1591-1641, 1635-1685, 1714-1764, and 1936-1986, respectively.

The present invention also relates to polypeptides of OTB0949, e.g., an isolated human OTB0949 polypeptide comprising or having the amino acid sequence set forth in SEQ ID NO 2, an isolated human OTB0949 polypeptide comprising an amino acid sequence having 80, 85, 90, 95, 97, 99% or more amino acid sequence identity to the amino acid sequence set forth in SEQ ID NO 1, and optionally having one or more of OTB0949 activities, such as cell signaling activity, secretory pathway activity, etc. Fragments specific to OTB0949 can also be used, e.g., to produce antibodies or other immune responses, as competitors to any of its activities, etc. These fragments can be referred to as being "specific for" OTB0949. The latter phrase, as already defined, indicates that the peptides are characteristic of OTB0949, and that the defined sequences are substantially absent from all other protein types. Such polypeptides can be of any size which is necessary to confer specificity, e.g., 5, 8, 10, 12, 15, 20, etc.

OTB182

OTB182 is an integral membrane protein comprising 307 amino acids (SEQ ID NO 12). Using SMART (e.g., Schultz et al., *Proc. Natl. Acad. Sci.*, 95:5857-5864, 1998; Letunic et al., *Nucleic Acid Res.*, 30:242-244, 2002), the protein is predicted to have seven membrane spanning regions at about amino acid positions 45-67, 87-109, 129-151, 161-183, 196-218, 231-253, and 278-297. There is a putative signal sequence (at amino acids 1-26) at its N-terminus which also overlaps with an eight transmembrane spanning domain (at amino acids

10-32).

OTB182 is expressed predominantly in excitable tissues, e.g., brain, heart, and muscle, with very low expression observed in prostate tissues. See Fig. 2. In the brain, it is expressed predominantly in thalamus. It also expressed in neural stem cells. Its expression
5 in excitable tissues makes OTB182 a highly selective marker for excitable tissues, as well as indicating a functional and/or developmental role in this tissue type. There are a number of genes whose expression is restricted predominantly to excitable cells, e.g., GABA-interacting factor-1 (GRIF-1; Beck et al., *J. Biol. Chem.*, 28 May 2002); calcium calmodulin dependent (CaM) kinase (e.g., Loseth et al., *Brain Research*, 869(1-2):137-145, 2000); sodium channel
10 types (e.g., Schaller et al., *J. Neurosci*, 12(4):1370-81, 1992); calcium dependent mitochondrial solute carrier (e.g., Del Arco et al., *J. Biol. Chem.*, 273(36):23327-23334, 1998).

A mouse homolog of OTB182 is AK003645 (SEQ ID NO 15) that codes for a 153 amino acid polypeptide. It shares about 92% sequence identity from amino acids 1-122, and
15 sharply diverges from that point onward. See Fig. 3. Murine OTB182 maps to chromosomal location 11E2. The present invention relates all transcripts associated with the AK003645 gene loci. The degree of nucleotide sequence identity between human and mouse (AK003645) is about 83% from about nucleotide position 71-437 of SEQ ID NO 12.

OTB182 is located at chromosomal band 17q25. A genetically-inherited
20 neuromuscular disease, hereditary neuralgic amyotrophy (HNA) has been mapped to this locus. See, e.g., Jeannet et al., *Neurology*, 57:1963-1968, 2001. In addition, a hereditary hearing loss maps to 17q25 (e.g., DFNA20, Morell et al., *Genomics*, 63:1-6, 2000) and mental retardation (Rio et al., *Human Genetics*, 108:511-515, 2001). Examples of specific polynucleotides are SEQ ID NOS 13 and 14.

25 Diseases or disorders which can be treated in accordance with the present invention include, but are not limited to neuropathy, neuralgic amyotrophy (e.g., HNA), myopathy, sensorineural hearing loss (e.g., DFNA20), mental retardation, neuromuscular disorders, brain cancer, such as a neuroblastoma, and other diseases and conditions involving heart, brain, and muscle tissues, etc.

30 A transgenic animal with OTB182 functionally disrupted can show a defect in an excitable cell. Such defect, includes, e.g., developmental defects, defects in the functional

activity of the cell, e.g., in excitability, membrane conductance, response to stimuli, signal transduction, or any of the disorders mentioned herein.

Antibodies to OTB182 can also be produced, e.g., an antibody which is specific-for: an epitope selected from amino acids 123-307 of SEQ ID NO 12, or comprising amino acid
5 27, 47, 64, 66, 75, 78, 105, 111, or 113 of SEQ ID NO 12

Human transient receptor potential cation channel (TRPCC) gene and polypeptide

Human TRPCC codes for a polypeptide of 1707 amino acids. As shown in Fig. 5, it is selectively expressed in brain, kidney, and pituitary, with very low expression observed in
10 testis and ovary. By the phrase "selectively expressed," it is meant that a nucleic acid molecule, when produced as a transcript, is characteristic of the tissue or cell-type in which it is made. This can mean that the transcript is expressed only in that tissue and in no other tissue-type, or it can mean that the transcript is expressed preferentially, differentially, predominantly, and more abundantly (e.g., at least 5-fold, 10-fold, etc., or more) in that tissue
15 when compared to other tissue-types.

The nucleotide and amino acid sequences of human TRPCC are shown in SEQ ID NOS 16 and 17. Analysis of its primary structure indicates the presence of six transmembrane domains at about amino acids 870-892, 901-1112, 904-921, 936-958, 971-990, 1005-1024, 1085-1107 of SEQ ID NO 17, however, by analogy to other ion channels, it
20 is generally believed to have only six transmembrane spanning regions. See, e.g., Clapham et al., *Nature Reviews, Neuroscience*, 2:387, 2001. The ion transport domain comprises amino acids 901-1112. There is also a putative transmembrane domain at the N-terminus at about amino acids 5-24. According to the six-transmembrane model, both the N- and C-terminus of the protein are intracellular, and provide a scaffolding for interaction with other proteins.

25 The human TRPCC contains 25 exons. The present invention relates to any isolated introns and exons that are present in the gene. Intron and exon boundaries can be routinely determined, e.g., using the sequences disclosed herein.

Partial sequences for human TRPCC were previously identified (e.g., Accession Numbers AB046836 and XM_036123). For example, human AB046836 (SEQ ID 19) is
30 incomplete, coding for 1017 amino acids (See Fig. 4, AB046836), and lacks the first 690 amino acids of human TRPCC, but shares about 99% identity with TRPCC along the rest of

-7-

its length. Another partial sequence, human XM_036123 (SEQ ID NO 18) codes for 988 amino acids (See Fig 4, XM_036123), lacking the first 719 amino acids of human TRPCC, but shares 100% identity with TRPCC along the rest of its length (See Fig 4). XM_140575 (SEQ ID NO 20) appears to be a homolog of human TRPCC, and shares about 94% sequence identity from about amino acids 82-693, or about amino acids 345-956 of human TRPCC (SEQ ID NO 17). Amino acids 1-81 and 694-736 (see Fig. 4) of the mouse homolog have low sequence identity with human TRPCC. Alignment with mouse genomic DNA using Spidey (NCBI) indicates that amino acids 1-80 of XM_140575 are derived from exons 1 and 2 of the genomic DNA, and amino acids 694-736 are derived from exon 7 of the mouse genomic DNA. XM_140575 is located on mouse chromosome 19B.

TRPCC maps to chromosomal region 9q21.1. Strikingly, hypomagnesemia with hypocalcemia (OMIM 602014) are known to be determined by a mutation within 9q21 (Walder et al., *Human Molecular Genetics*, 6: 1491-1497, 1997), as would be expected with a channel responsible for cation conductance. Consistent with its expression in brain, a susceptibility to amyotrophic lateral sclerosis with frontotemporal dementia (OMIM 105550) was mapped to this same chromosomal locus (Pinsky et al., *Clinical Genetics*, 7:186-191, 1975; Hosler et al., *JAMA*, 284:1664-1669, 2000). In addition, schizophrenia (Hovatta et al., *Am. J. Hum. Genet.*, 65:1114-24), and familial dyskinesia/facial myokymia (Fernandez et al., *Ann. Neurol.*, 49:486-92, 2001) are also associated with this gene locus. Nucleic acids of the present invention can be used, e.g., as linkage markers, diagnostic targets, and therapeutic targets for any of the mentioned disorders, as well as any disorders or genes mapping in proximity of TRPCC.

TRCC polynucleotides, polypeptides, ligands, and binding partners thereto, can be used in a number of useful ways. For example, binding partners, such as antibodies and ligands, can be used to selectively target agents to brain, kidney, and other tissues in which it is expressed for purposes including, but not limited to, imaging, diagnostic, therapeutics, etc. Imaging of tissues can be facilitated using agents such as TRPCC antibodies that can be used to target contrast agents to a specific site in the body. Various imaging techniques have been used in this context, including, e.g., X-ray, CT, CAT, MRI, ultrasound, PET, SPECT, and scintigraphic. A reporter agent can be conjugated or associated routinely with a TRPCC antibody. Ultrasound contrast agents combined with ligands such as antibodies are described

in, e.g., U.S. Pat. Nos 6,264,917; 6,254,852; 6,245,318; and 6,139,819. MRI contrast agents, such as metal chelators, radionucleotides, paramagnetic ions, etc., combined with selective targeting agents are also described in the literature, e.g., in U.S. Pat. Nos. 6,280,706 and 6,221,334. The methods described therein can be used generally to associate TRPCC and
5 ligands thereof with an agent for any desired purpose.

An active agent can be associated in any manner with an TRPCC ligand that is effective to achieve its delivery to the target. The association of the active agent and the ligand ("coupling") can be direct, e.g., through chemical bonds between the binding ligand and the agent or via a linking agent, or the association can be less direct, e.g., where the
10 active agent is in a liposome, or other carrier, and the ligand is associated with the liposome surface. In such case, the ligand can be oriented in such a way that it is able to bind to TRPCC on the surfaces of kidney or brain cells.

Useful human TRPCC polypeptides and corresponding nucleic acids include polypeptides comprising amino acids 1-88, 5-24, 1-690, 1-719, and fragments thereof (See
15 SEQ ID NO 17 and Fig. 4). Nucleic acids and polypeptides can be used as probes (e.g., in PCR, in Northern blots, etc.), as diagnostic agents, to generate antibodies, as vaccines, to produce recombinant proteins, as antisense, etc. A specific polynucleotide according to the present invention can be determined routinely. Examples are specific probes are SEQ ID NOS 21-24, e.g., where SEQ ID NOS 23 and 24 can be used as forward and reverse PCR
20 primers, respectively, to amplify a portion of amino acid region 1-160 of SEQ ID NO 17.

TRPCC has a number of biological activities, including, e.g., cation transport, signal transduction, protein binding, etc. By "signal transduction" is meant the activation of a chain of events that alters the concentration of one or more small intracellular signaling molecules (second messengers), e.g., cyclic AMP, calcium ions, as described in Sabala et al.,
25 *British Journal of Pharmacology*, 132:393-402, 2001. By "cation transport" is meant the influx or efflux of a cation, e.g., calcium, magnesium, into or from a cell. Mizuno et al., *Molecular Brain Research*, 64:41-51, 1999. Protein binding indicates the ability of the protein to interact with other proteins, e.g., as the N-terminus interacts with intracellular proteins. These activities can be determined routinely. Signal transduction can be assessed
30 by expression of TRPCC in cells, etc., and measurement of the concentrations of elicited second messengers or byproducts, e.g., Ca^{2+} or Mg^{2+} or cAMP, inositol, etc., by, e.g., atomic

absorption spectrometry (ThermoElemental SOLAAR AA spectrometers), radioimmunoassay, etc. Sano et al. *Science*, 293:1327-1330, 2001. Cation transport can be assessed by measurement of changes in ionic currents by whole-cell patch-clamp analysis. For instance, cells or oocytes can be transfected with a polynucleotide of the present invention and then analyzed for expression of calcium channel activity, e.g., using patch clamp, calcium activated dyes, etc.. See, also, e.g., Strubing et al., *Neuron*, 29:645-655, 2001; Sano et al., *Science*, 293:1327, 2001; Ohki et al., *J. Biol. Chem.*, 275:39055-39060, 2000; Boulay et al., *J. Biol. Chem.*, 272:29672-29680, 1997.

The present invention relates to an isolated polynucleotide comprising, e.g., a polynucleotide sequence coding without interruption for a human TRPCC polypeptide, or complement thereto, said TRPCC having 80%, 85%, 90%, 92%, 95%, 99%, or more amino acid sequence identity along its entire length to the sequence comprising amino acids 1-690 of SEQ ID NO 17, and 80%, 85%, 90%, 92%, 95%, 99%, or more amino acid sequence identity along its entire length to the sequence comprising from amino acids 691-1707 of SEQ ID NO 17, and which has, e.g., cation transport, signal transduction, or protein binding activity.

Antibodies can be prepared against specific epitopes or domains of TRPCC, e.g., amino acids 2-30, 773-789, 870-887, 905-913, 943-958, 969-986, 1005-1022, 1087-1114, 1125-1131, 789-870, 913-943, 986-1005, etc.

Detection can be desirable for a variety of different purposes, including research, diagnostic, prognostic, and forensic. Diagnostic purposes included testing patients and their families for the presence of mutations associated with hypomagnesemia with hypocalcemia or amyotrophic lateral sclerosis with frontotemporal dementia. The selected mutant alleles, SNPs, polymorphisms, etc., can be used diagnostically to determine whether a subject has, or is susceptible to a disorder associated with TRPCC, as well as to design therapies and predict the outcome of the disorder. Methods involve, e.g., diagnosing a disorder associated with TRPCC or determining susceptibility to a disorder, e.g., hypomagnesemia with hypocalcemia or amyotrophic lateral sclerosis with frontotemporal dementia, comprising, detecting the presence of a mutation in a TRPCC gene (such as a mutation in SEQ ID NO 16, or variants thereof. The sequences of TRPCC genes can also be compared, e.g., between a normal gene as shown in SEQ ID NO 16 and the sequence of a gene from a patient with the disorder, e.g.,

hypomagnesemia with hypocalcemia.

Fragments specific to TRPCC can also be used, e.g., to produce antibodies or other immune responses, as competitors to nucleotide binding, ligand binding, etc. or as, e.g., inhibitors or stimuli in signal transduction pathways. These fragments can be referred to as being "specific for" TRPCC. The latter phrase, as already defined, indicates that the peptides are characteristic of TRPCC, and that the defined sequences are substantially absent from all other protein types. Such polypeptides can be of any size necessary to confer specificity, e.g., 5, 8, 10, 12, 15, 20, etc. Examples of polypeptides include but are not limited to polypeptides that comprise the following amino acid residues: 2-60, 598-660 of SEQ ID NO 17, or fragments thereof.

Biological activities of TRPCC include, e.g., cation channel activity, signal transduction activity, and protein binding activity. As discussed above, the biological activity of TRPCC can be measured routinely. For example, if agents are to be identified which modulate the channel activity of TRPCC either electrophysiology or calcium imaging can be used to assess their effects, e.g., using fluo-3, Fura-red, Ca-sensitive chemi-luminescent probes, etc. (e.g., kits are commercially available from Molecular Probes) and a laser scanning confocal microscope to visualize the changes in intracellular calcium as a result of modulation of TRPCC.

A transgenic animal, or animal cell, lacking one or more functional TRPCC genes can be useful in a variety of applications, including, as an animal model for hypomagnesemia with secondary hypocalcemia, amyotrophic lateral sclerosis with frontotemporal dementia, etc., drug screening assays (e.g., for signal transduction mediated by agents other than TRPCC; by making a cell deficient in TRPCC, the contribution of other receptors to, e.g., Ca^{2+} modulation can be specifically examined), as a source of tissues deficient in TRPCC activity, etc. Such an animal can show a defect in cation (e.g., calcium) conductance, e.g., an impairment in the permeation of an ion through the channel.

An isolated polynucleotide can comprise, e.g., a polynucleotide sequence coding without interruption for a human TRPCC polypeptide, or complement thereto, said TRPCC having 90% or more amino acid sequence identity along its entire length to the sequence comprising amino acids 1-690 of SEQ ID NO 17, and 90% or more amino acid sequence identity along its entire length to the sequence comprising from amino acids 691-1707 of

-11-

SEQ ID NO 17, and which has cation transport activity.

The present invention also relates to a methods of identifying a mutation associated with amyotrophic lateral sclerosis with frontotemporal dementia, comprising: comparing the structure of: genomic DNA comprising all or part of human TRPCC, mRNA comprising all or part of human TRPCC, cDNA comprising all or part of human TRPCC, or a polypeptide comprising all or part of human TRPCC, with the complete structure of human TRPCC as set forth in SEQ ID NO 16, in a patient having amyotrophic lateral sclerosis with frontotemporal dementia, or a family member thereof.

10 Melanocortin

The present invention relates to novel forms of a melanocortin-1 receptor (also known as "MCR-1" or alpha-melanocyte stimulating hormone receptor). It is highly expressed in melanocytes, and is a key component of the pathway which modulates skin and hair pigmentation. Moreover, certain alleles of MCR-1 are associated with a high risk of melanoma. MCR-1 is also expressed in other tissues, including monocytes, mast cells, placenta, pituitary, and endothelial cells.

MCR-1 belongs to the G-protein coupled receptor (GPCR) super-family. Its expression is restricted to melanocytes and few other cell types, such as monocytes, mast cells, and endothelial cells. See, e.g., Smith et al., *Gene*, 281:81-94, 2001; Scholzen et al., *Annals of the New York Academy of Sciences*, 885:239-253 (1999). Stimulation of the receptor by its natural ligands (e.g., alpha-melanocyte stimulating hormone or "α-MSH") causes an increase in cAMP levels which, in turn, stimulates intracellular tyrosinase activity. Increased activity of the tyrosinase enzyme drives the conversion of pheomelanin (yellow and red pigments) to eumelanin (brown and black pigments).

The MCR-1 gene is located at chromosomal position 16q24. It is adjacent to the tubulin TUBB4 gene, and its 3' region overlaps with the tubulin promoter (Smith et al.). Transcripts containing genic material from both MCR1 and TUBB4 have been identified, including transcripts which contain coding sequences from both. See, e.g., NCBI accession number BC020171. These may be involved in cancer.

Almost 40 different polymorphisms in the MCR receptor have been identified. See, Sturm et al., *Gene*, 277:49-62, 2001; Table 1. Several of these (e.g., Arg151Cys; Arg160Trp;

-12-

Asp294His) are strongly associated with red hair, fair skin, and poor tanning ability. It has been reported that these alleles are nonfunctional receptors and do not stimulate cAMP production when stimulated by MSH. See, Table 2. As a consequence, pheomelanin is not converted to eumelanin, and skin and hair color reflect the cell's high content of the yellow and red pheomelanin pigments. Significantly, individuals who have these alleles are also at a higher risk for skin cancers, such as basal cell carcinoma, squamous cell carcinoma, and melanoma. See, e.g., Sturm et al., *Am. J. Hum. Genet.*, 6 (supplement to volume 67): 16, Oct. 2000. See, also OMIM, No. 155555 for other information on MCR-1, including disease information, polymorphisms, etc.

The present invention relates to novel forms of MCR-1. In one embodiment, the present invention relates to a novel MCR-1 variant, MCR-1 type C or MCR-1C, which possesses a unique carboxy-terminus. Previous reports had identified a 317 amino acid form of MCR-1 ("MCR-1A") in a number of different species, including human (SEQ ID NO 30), chimpanzee, muskox, sheep, cow, horse, dog, and fox. This form was characterized as full-length. Several minor size variants were observed, as well, e.g., in mouse (315 amino acid acids), in pig (310 amino acids), and in cow (321 amino acids). A second form, MCR-1B, was (SEQ ID NO 31) also reported that had an additional 65 amino acids at its terminus (Tan et al., *FEBS Letters*, 451:137-141, 1991; WO 00/39147). The present invention relates to a third form of MCR-1 (MCR-1C) that comprises 32 carboxy-amino acids (amino acids 367-398 of SEQ ID NO 26) not previously identified in any melanocortin receptor variant. This novel form comprises part of the new carboxy terminus identified in MCR-1B, but diverges from it at amino acid position 367. See, Fig. 6.

Fig. 7 shows exons which have been detected in melanocortin-1 receptors. Exons 1, 2, and 3 contain MCR coding sequences; Exons 5, 6, and 7 contain tubulin coding sequences.

MCR-1A (e.g., NM_002386: SEQ ID NO 30) contains exon 1, and MCR-1B contains exons 1 and 2. MCR-1C contains coding sequence from exons 1-3. As indicated by the stop codon TGA, exon 3 comprises both coding and noncoding sequence. MCR-1C can also contain noncoding sequences, e.g., exons 4, 5, 6, and/or 7 (e.g., for a total of exons 1-7). BC020171, mentioned above, contains the coding sequence from exons 1 and 2 fused to the coding sequences of tubulin in exons 5-7. Examples of promoters for MCR-1C include, e.g., SEQ ID NOS 35-37.

-13-

The present invention also relates to a polymorphism at amino acid position 120, where an isoleucine (I) is replaced with a threonine (T). Isoleucine is present at amino acid position 120 in most melanocortin receptor-1 homologs, except pig which has a methionine substitution. This polymorphism may affect the receptor's functionality. Analysis of the transmembrane structure using TMHMM v. 2.0 (Krogh et al., *Journal of Molecular Biology*, 305(3):567-580, January 2001; Sonnhammer et al., In J. Glasgow et al., editors, *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology*, pages 175-182, Menlo Park, CA, 1998. AAAI Press; Moller et al., *Bioinformatics*, 17(7):646-653, July 2001) reveals a different number of predicted transmembrane sequences than the isoleucine isoform.

The present invention relates to any polynucleotide, or polypeptide encoded thereby, which codes for MCR-1C, including receptors having any polymorphism, such as the naturally-occurring polymorphisms listed in Tables 1 and 2, and those disclosed herein. Examples include SEQ ID 27 having 120T and 163Q, SEQ 28 having 120I and 163R, and SEQ ID 29 having 120I and 163Q. It also includes polynucleotide and polypeptide fragments which are specific for MCR-1C (e.g., 367-398 of SEQ ID NO 26 and fragments thereof), and polynucleotides and polypeptides which comprise such specific fragments. For example, the present invention relates to a polynucleotide comprising a coding sequence for amino acid 367-398 of SEQ ID NO 26, or fragments thereof, such as any five amino acid sequence contained therein.

The present invention also relates to an isolated polynucleotide comprising, a polynucleotide sequence coding without interruption for a human MCR-1C, said MCR-1C having about 80%, 85%, 88%, 89%, 90%, 92%, 95%, 99%, etc., or more amino acid sequence identity along its entire length to the amino acid sequence set forth in SEQ ID NO 26, or a complement thereto, and which has ligand-binding activity, G-protein binding activity, or cAMP production activity. For example, such a polynucleotide can comprise one or more of the polymorphisms listed in Tables 1 and 2 (e.g., if 36 of the listed polymorphisms were present in such a polynucleotide, it would have about 90% (360/398) sequence identity along its entire length to the amino acid sequence of SEQ ID NO 26. A corresponding amount of nucleotide is included, e.g., 90%, 92%, 95%, 97%, 98%, 99%, or more.

Similarly, the present invention relates to an isolated polynucleotide comprising, a polynucleotide sequence coding without interruption for a human MCR-1C, or complement thereto, said MCR-1C having 80%, 84%, 85%, 86%, 88%, 90%, 95%, or more amino acid sequence identity along its entire length to the sequence comprising amino acids 1-316 of SEQ ID NO 26, and 85%, 90%, 92%, 95%, etc. or more amino acid sequence identity along
5 its entire length to the sequence comprising from amino acids 317-398 of SEQ ID NO 26, and which has ligand-binding activity, G-protein binding activity, or cAMP production activity.

As stated, a polynucleotide can code for a polypeptide having one or more of the
10 following activities, ligand-binding activity, G-protein binding activity, cAMP production activity, or other functional activities. Ligand-binding activity indicates the ability of MCR-1C to bind specifically to a receptor ligand, such as a ACTH, MSH, etc. Ligand binding activity can be using a radioactive or otherwise labeled ligand, or whole-cell assays using labeled ligands. See, e.g., WO0039147, WO9957148, U.S. Pat. Nos. 5,731,408, 6,100,048,
15 and 6,350,760; Libert et al., *Pigment Cell Res.*, 2:510-518, 1989.

G-protein binding activity indicates the ability of the receptor protein to bind to a G-protein. Such binding does not determined routinely, e.g., using filtration assays necessarily have to be productive, i.e., the binding does not have to result in stimulation of the signal transduction cascade. G-protein binding can be measured using in vivo and in vitro binding
20 assays, as well as functional assays. See, e.g., Ford et al., *Science*, 280:1271-1274, 1998.

cAMP production is a measure of the ability of the receptor to stimulate the generation of cAMP upon binding by a receptor agonist. MCR is known to couple to G-proteins and thereby activate adenylyl cyclase, increasing intracellular levels of cAMP (e.g., Buckley & Ramachandran, *Proc. Natl. Acad. Sci.*, 78: 7431-7435, 1981; Grahame-Smith et al., 1967, *J. Biol. Chem.* 242: 5535-5541; Mertz & Catt, 1991, *Proc. Natl. Acad. Sci.* 88: 8525-8529; Pawalek et al., 1976, *Invest. Dermatol.* 66: 200-209). This property of cells
25 expressing the MCR-1C can be used assess its "cAMP production activity." For example, cells can be transfected with MCR-1C DNA, plated, and washed once with DMEM containing 1% bovine serum albumin (BSA) and 0.5 mM IBMX (a phosphodiesterase
30 inhibitor). The cells can then be treated with hormone (e.g., alpha-MSH, gamma-MSH, ACTH, etc.). Following hormone treatment, the cells can be washed with phosphate buffered

-15-

saline, and intracellular cAMP extracted by lysing the cells. Intracellular cAMP concentrations can be determined routinely, e.g., using an assay (Amersham) which measures the ability of cAMP to displace cAMP from a high affinity cAMP binding protein (see Gilman, 1970, *Proc. Natl. Acad. Sci.*, 67: 305-312).

5 Polynucleotide and polypeptides of the present invention can be used for a variety of purposes, including, but not limited to, treating cancers, treating skin cancer and other cancers modulating skin and hair pigmentation, identifying MCR ligands, modulating the MCR-1 receptor types, determining susceptibility to skin cancer, detecting MCR-1C expression, determining polymorphisms in MCR-1C, making MCR-1C polypeptide,
10 expressing MCR-1C in host cells, making antibodies to MCR-1 receptor types, modulating cutaneous inflammation (see, e.g., Bhardwaj et al., *J. Immunol.*, 158:3378-3384, 1997; Luger et al., *Ann. NY Acad. Sci.*, 917:232-238, 2000), modulating melanocytes, monocytes, endothelial cells, or other cells in which MCR-1C is expressed, etc.

The expression of MCR-1C on the surface of melanoma cells makes it a useful target.
15 Melanoma is a skin cancer which originates from melanocytes present normally in the epidermis and underlying cell layers. There are four basic types: lentigo maligna melanoma, superficial spreading melanoma, nodular melanoma, and acral lentiginous melanoma. Because of its expression on melanocytes, MCR-1C specific antibodies and other binding partners can be used to treat melanoma, e.g., by conjugating cytotoxic agents to antibodies directed to the
20 receptor. In addition, MCR-1C polynucleotides, polypeptides, and binding partners thereto can be used to detect metastatic melanoma cells.

Modulation of the MCR-1C can also be used to modulate skin pigmentation, e.g., to increase the amount brown and black pigments to darken skin color, to provide protective effects against UV radiation, to block receptor activation, e.g., preventing or reducing the
25 accumulation of brown and black pigments in the skin, preventing or reducing tanning, preventing or reducing skin freckling, etc. Agonists and antagonists of the melanocortin receptor, include, alpha-melanocyte stimulating hormone and adrenocorticotrophic hormone. Other ligands are disclosed in, e.g., WO9957148, U.S. Pat. Nos. 5,731,408, 6,100,048, and 6,350,760, and can be identified and isolated as described in these patents, as well as
30 WO0039147.

As discussed earlier, several MCR-1 alleles have been associated with a greater risk

-16-

of skin cancer. For example, the presence of the Asp84Glu variant imposed a high risk of melanoma in individual carriers. See, Kennedy et al., *J. Invest. Dermatol.*, 117:294-300, 2001. Other alleles with increased risk of melanoma included, Val60Leu, Val92Met, Arg142His, Arg151Cys, Arg160Trp, Arg163Gln, and His260Pro (Kennedy et al.). See, also, Scott et al., *J. Cell. Sci.*, 115 (Pt. 11):2349-2355, 2002. MCR-1C of the present invention can be used to assess melanoma risks, e.g., determining the presence of a variant of MCR-1C in individuals, and whether such variants are associated with skin cancer and other melanocyte disorders. Analysis can be performed by any suitable method, e.g., by single-stranded conformation polymorphism analysis and DNA sequence analysis.

Expression can also be "selective," where expression is observed. By the phrase "selectively expressed," it is meant that a nucleic acid molecule comprising the defined sequence of nucleotides, when produced as a transcript, is characteristic of the tissue or cell-type in which it is made. This can mean that the transcript is expressed only in that tissue and in no other tissue-type, or it can mean that the transcript is expressed preferentially, differentially, and more abundantly (e.g., at least 5-fold, 10-fold, etc., or more) in that tissue when compared to other tissue-types.

In view of their selectivity and display on the cell surface, MCR-1C polypeptides of the present invention are a useful target for histological, diagnostic, and therapeutic applications relating to the cells in which they are expressed. Antibodies and other protein binding partners (e.g., ligands, aptamers, small peptides, etc.) can be used to selectively target agents to a tissue for any purpose, included, but not limited to, imaging, therapeutic, diagnostic, drug delivery, gene therapy, etc. For example, binding partners, such as antibodies, can be used to treat melanomas in analogy to how c-erbB-2 antibodies are used to breast cancer. They can also be used to detect metastatic cells, in biopsies, etc. The genes and polypeptides encoded thereby can also be used in tissue engineering to identify tissues as they appear during the differentiation process, to target tissues, to modulate tissue growth (e.g., from starting stem cell populations), etc. Useful antibodies or other binding partners include those that are specific for parts of the polypeptide which are exposed extracellularly. Any of the methods described above and below can be accomplished in vivo, in vitro, or ex vivo.

Binding partners can also be used as to specifically deliver therapeutic agents to a

-17-

tissue of interest. For example, a gene to be delivered to a tissue can be conjugated to a binding partner (directly or through a polymer, etc.), in liposomes comprising cell surface, and then administered as appropriate to the subject who is to be treated. Additionally, cytotoxic, cytostatic, and other therapeutic agents can be delivered specifically to the tissue to
5 treat and/or prevent any of the conditions associated with the tissue of interest.

The present invention relates to methods of detecting melanoma cells, comprising one or more of the following steps, e.g., contacting a sample comprising cells with a polynucleotide specific for MCR-1C (e.g., amino acids 367-398, and fragments thereof), or a mammalian homolog thereof, under conditions effective for said polynucleotide to hybridize
10 specifically to said gene, and detecting specific hybridization. Detecting can be accomplished by any suitable method and technology, including, e.g., any of those mentioned and discussed below, such as Northern blot and PCR. Specific polynucleotides include SEQ ID NOS 32-34, and complements thereto.

As indicated above, binding partners can be used to deliver agents specifically to
15 melanocytes, e.g., for diagnostic, therapeutic, and prognostic purposes, including the treatment of melanoma. Methods of delivering an agent to a melanocyte cell can comprise, e.g., contacting a melanocyte with an agent coupled to binding partner specific for a melanocortin receptor gene of the present invention, whereby said agent is delivered to said cell. Any type of agent can be used, including, therapeutic and imaging agents. Contact with
20 the melanocyte (e.g., a melanoma) can be achieved in any effective manner, including by administering effective amounts of the agent to a host orally, parentally, locally, systemically, intravenously, etc. The phrase "an agent coupled to binding partner" indicates that the agent is associated with the binding partner in such a manner that it can be carried specifically to the target site. Coupling includes, chemical bonding, covalent bonding, noncovalent bonding
25 (where such bonding is sufficient to carry the agent to the target), present in a liposome or in a lipid membrane, associated with a carrier, such as a polymeric carrier, etc. The agent can be directly linked to the binding partner, or via chemical linkers or spacers.

Imaging of specific organs can be facilitated using tissue selective antibodies and other binding partners that selectively target contrast agents to a specific site in the body.

30 Various imaging techniques have been used in this context, including, e.g., X-ray, CT, CAT, MRI, ultrasound, PET, SPECT, and scintigraphic. A reporter agent can be conjugated or

-18-

associated routinely with a binding partner. Ultrasound contrast agents combined with binding partners, such as antibodies, are described in, e.g., U.S. Pat. Nos. 6,264,917, 6,254,852, 6,245,318, and 6,139,819. MRI contrast agents, such as metal chelators, radionucleotides, paramagnetic ions, etc., combined with selective targeting agents are also described in the literature, e.g., in U.S. Pat. Nos. 6,280,706 and 6,221,334. The methods described therein can be used generally to associate a partner with an agent for any desired purpose.

A transgenic animal with a disrupted melanocortin-1C receptor can have a pigmentation phenotype, e.g., red or fair hair. Functional disruption of the gene can be accomplished in any effective way, including, e.g., introduction of a stop codon into any part of the coding sequence, e.g., to prevent expression of amino acids 367-398, such that the resulting polypeptide is biologically inactive or lacks one or more of its functional regions, introduction of a mutation into a promoter or other regulatory sequence that is effective to turn it off, or reduce transcription of the gene, insertion of an exogenous sequence into the gene which inactivates it (e.g., which disrupts the production of a biologically-active polypeptide or which disrupts the promoter or other transcriptional machinery), deletion of sequences from the gene, etc. A transgenic animal, or animal cell, lacking one or more functional genes of the present invention can be useful in a variety of applications, including, as an animal model for conditions and diseases associated with melanocortin-1C, for drug screening (e.g., by making a cell deficient in MCR-1C, the contribution of the activity remaining variants, such as MCR-1B and the 317-amino acid form, can be assessed), as a source of tissues deficient in one or more MCR-1 activities. The animal's endogenous locus can be replaced with a continuous coding sequence for MCR-1C, such that only MCR-1C is expressed, and no other form, such as MCR-1B and the 317-amino acid form, are expressed.

PCR based methods can also be used in the methods of detecting polynucleotides for human MCR-1C. In such methods, more than one probe specific for MCR-1C can be used, e.g., a pair of specific polynucleotide probes which are capable of amplifying a polynucleotide sequence of MCR-1C, such as corresponding to amino acids 1-366, 367-398, etc., of SEQ ID NO 26. For instance, SEQ ID NO 32 is in exon 1, SEQ ID NO 33 spans exons 2-3, and SEQ ID NO 34 is in exon 4. Thus, in a PCR reaction, SEQ IDS 32 and 33 produce a fragment about 262 base pairs that is absent in MCR-1A and MCR-1B. SEQ ID

NOS 32 and 34 in a PCR reaction produce a fragment of about 615 base pairs which is absent from MCR-1A and MCR-1B.

Mutant alleles, polymorphisms, SNPs, etc., can be identified and isolated from melanomas and other skin conditions that are known, or suspected to have, a genetic component. Identification of such genes can be carried out routinely (see, above for more guidance), e.g., using PCR, hybridization techniques, direct sequencing, mismatch reactions (see, e.g., above), RFLP analysis, SSCP (e.g., Orita et al., *Proc. Natl. Acad. Sci.*, 86:2766, 1992), etc., where a polynucleotide having a sequence selected from SEQ ID NO 25 (especially corresponding to amino acids 367-398) can be used as a probe. The selected mutant alleles, SNPs, polymorphisms, etc., can be used diagnostically to determine whether a subject has, or is susceptible to a melanoma or other condition (e.g., pigmentation variation, inflammatory condition) associated with a melanocortin receptor gene of the present invention, as well as to design therapies and predict the outcome of the disorder.

The present invention an isolated polynucleotide comprising, a polynucleotide sequence coding without interruption for a human MCR-1C, or complement thereto, said MCR-1C having 84% or more amino acid sequence identity along its entire length to the sequence comprising amino acids 1-316 of SEQ ID NO 26, and 90% or more amino acid sequence identity along its entire length to the sequence comprising from amino acids 317-398 of SEQ ID NO 26, and which has ligand-binding activity, G-protein binding activity, or cAMP production activity.

OTB860

OTB860 codes for an intracellular polypeptide comprising 1700 amino acids. Its nucleotide and amino acid sequences are shown in SEQ ID NOS 38 and 39. Expression of OTB860 is detected predominantly in heart and brain tissues, with minimally detectable levels in breast and testes tissues. The expression pattern is illustrated in Fig. 8.

As shown in Fig. 9, OTB860 is related to KIAA1678 (also known as AB051465; SEQ ID NO 42). It contains 369 amino acids at its N-terminal end which are not present in KIAA1678, and a 29 amino acid insertion (1546-1574) at about amino acid position 1545. It also differs at amino acid positions 847 (histidine instead of glutamine) and 867 (arginine instead of glutamine).

Consistent with the expression pattern of OTB860, there are a number of functional and developmental pathways that are shared between neuronal and cardiac cells. For example, neuregulins (neuregulin-1 – NRG1) play an important role in myocardial and neuronal development. Mice deficient in IgL-domain containing neuregulins have severe defects in the developing heart and nervous system. See, e.g., Kramer et al., *Proc. Natl. Acad. Sci.*, 93: 4833-4838, 1996; Zhao et al., *J. Biol. Chem.*, 273:10261-10269, 1998. These effects appear to be mediated by Type I neuregulins. Meyer et al., *Develop*, 124:3575-3586, 1997. The neuroregulin receptor, erbB4, is also highly expressed in heart and brain, although its expression is not restricted to these tissues (data not shown). Adhesion pathways are also shared between the two tissues. For example, alpha4-integrin is expressed in both neuronal and cardiac cells. See, e.g., Pinco et al., *Mech. Dev.*, 100:99-103, 2001.

OTB860 maps to chromosomal band 2q36. A number of disorders have been mapped to, or in close proximity to, this chromosome location. These include, e.g., brachydactyly type A1, pili torti and nerve deafness, syndactyly type 1, Gracile syndrome (growth retardation and early death), and epilepsy. Recently, in a whole genome scan, 2q36 was identified as a locus associated with acute coronary syndrome (Harrap et al., *Arterioscler. Thromb. Vasc. Biol.*, 22:874-878, 2002), involving myocardial infarction, unstable angina, atherosclerotic plaque disruption, and coronary thrombosis. Nucleic acids of the present invention can be used as linkage markers, diagnostic targets, therapeutic targets, for any of the mentioned disorders, as well as any disorders or genes mapping in proximity to it.

OTB860 can be used in diagnostic, therapeutic, prophylactic, and research applications. RNA and polypeptide detection methods can be used to determine whether a sample comprises neuronal or cardiac tissues. When a positive is obtained, cell type markers can be used to determine precisely whether the tissue is neuronal or cardiac. For example, the presence or absence of a neuronal marker would distinguish between brain and heart tissues. Non-limiting examples of neuronal markers include, presenilins, genes and polypeptides in the pathways for neurotransmitter synthesis, receptor, metabolism, etc., (e.g., serotonin, MAO, dopamine, norepinephrine, nitric oxide, etc.), apolipoprotein A, APP, neuron-specific enolase (NSE), glial fibrillary acidic protein (GFAP), S100, GAP-43, neuron-specific beta-III tubulin, Stac (neuron-specific protein with an SH3 domain, e.g., *Genomics*, 47:140-2, 1998), myelin basic protein, vimentin, etc. Non-limiting examples of heart tissue

markers include, cardiac troponin I, and smooth muscle markers such as CRP1.

The present invention also relates to polypeptide detection methods for assessing heart or brain function, e.g., methods of assessing heart or brain function, comprising, detecting OTB860 polypeptide, or fragments thereof, in a body fluid, whereby the level of
5 OTB860 polypeptide in said fluid is a measure of heart or brain function. Heart or brain function tests are usually performed to determine whether the organ is functioning normally as a way of diagnosing disease. Various tests are used to test for heart function, such as electrocardiogram, stress test, echocardiogram, oxygen levels, and cardiac enzyme tests (e.g., creatine phosphokinase, troponin, lactate dehydrogenase, and myoglobin). Detection of
10 OTB860 provides an additional assessment tool, especially in diseases such as myocardial infarction and other conditions, e.g., where cellular debris, etc., is released systemically. As with the other tests, elevated levels of OTB860 in blood, or other fluids, can indicate impaired brain or heart function. Values can be determined routinely, as they are for other functional markers.

15 OTB860 polynucleotides and polypeptides can be used to treat, prevent and diagnose diseases of the heart, including, e.g., acute coronary syndrome, myocardial hypertrophy, heart failure, conduction disorders, arrhythmias, bradyarrhythmias, sinus node dysfunction, tachyarrhythmias, tachycardias, atrial fibrillation, congenital heart diseases (see, e.g., Harrison's *Principles of Internal Medicine*, Volume 1, 12th Edition, 1991, Pages 924-925),
20 atrial and ventricular septal defects, congenital aortic stenosis, coarctation of the aorta, valvular heart disease, myocardial infarction, ischemic heart disease, cardiomyopathy, pericardial diseases, cardiac tumors (e.g., myxoma, lipoma, papillary fibroelastoma, rhabdomyoma, sarcoma, etc), coronary artery disease, atherosclerosis, aortic aneurysm, etc.

OTB860 polynucleotides and polypeptides can also be used to treat, prevent and
25 diagnose diseases of the brain, including, e.g., vascular diseases, hypoxia, ischemia, infarction, tumors, neuroglial tumors, astrocytoma, glioblastoma multiforme, pilocytic astrocytoma, oligodendroglioma, ependymoma, choroid plexus papilloma, neuronal tumors, neuroblastoma, ganglioneuroma, gangliocytoma, ganglioglioma, primitive or undifferentiated
30 tumors, medulloblastoma, tumors of meninges, meningioma, lymphomas, demyelinating diseases, multiple sclerosis, periventricular encephalitis, degenerative diseases, Alzheimer's, Pick's, Huntington's, Parkinsonism, ALS, Werdnig-Hoffman, degenerative diseases of

cerebral cortex, ganglia, brainstem, and motor neurons, inborn errors of metabolism, demyelinating and dysmyelinating disorders, Pelizaeus-Merzbacher disease, multiple sclerosis, various leukodystrophies, post-traumatic demyelination, cerebrovascular (CVS) accidents, neuritis, neuropathies, particularly, multifocal leucoencephalopathy, Guillain-Barre syndrome, retrobulbar neuritis, acute rubella encephalitis, chronic relapsing polyneuropathy, 5 intravascular lymphomatosis, Krabbe disease, globoid cell leukodystrophy, subacute combined degeneration of the spinal cord and brain, allergic encephalitis, murine coronavirus, hepatitis virus infection, or Theiler's murine encephalomyelitis, prion diseases, Creutzfeldt-Jakob, especially, febrile familial convulsions, epilepsy, vascular neuromyopathy, 10 cerebellar ataxia, etc.

When expression is described as being "predominantly" in a given tissue, this indicates that the gene's mRNAs levels are highest in this tissue as compared to the other tissues in which it was measured. Expression can also be "selective," where expression is observed. By the phrase "selectively expressed," it is meant that a nucleic acid molecule 15 comprising the defined sequence of nucleotides, when produced as a transcript, is characteristic of the tissue or cell-type in which it is made. This can mean that the transcript is expressed only in that tissue and in no other tissue-type, or it can mean that the transcript is expressed preferentially, differentially, and more abundantly (e.g., at least 5-fold, 10-fold, etc., or more) in that tissue when compared to other tissue-types.

20 The present invention relates to methods of detecting brain or heart cells, comprising one or more of the following steps, e.g., contacting a sample comprising cells with a polynucleotide specific for OTB860 (e.g., SEQ ID NOS 40-41), or a mammalian homolog thereof, under conditions effective for said polynucleotide to hybridize specifically to said gene, and detecting specific hybridization. Detecting can be accomplished by any suitable 25 method and technology, including, e.g., any of those mentioned and discussed below, such as Northern blot and PCR. Specific polynucleotides include SEQ ID NOS 40-41, and complements thereto.

Detection can also be achieved using binding partners, such as antibodies (e.g., monoclonal or polyclonal antibodies) that specifically recognize polypeptides coded for by 30 genes of the present invention. Thus, the present invention relates to methods of detecting a brain or heart cell, comprising, one or more the following steps, e.g. contacting a sample

-23-

comprising cells with a binding partner (e.g. an antibody, an Fab fragment, a single-chain antibody, an aptamer) specific for a polypeptide coded for by OTB860 (e.g., SEQ ID NO 39), or a mammalian homolog thereof, under conditions effective for said binding partner bind specifically to said polypeptide, and detecting specific binding. Protein binding assays can be accomplished routinely, e.g., using immunocytochemistry, ELISA format, Western blots, etc. Useful epitopes include those exposed to the surface.

A brain or heart cell (see above for examples of brain or heart cell types) can also be modulated in accordance with the present invention, e.g., by methods of modulating a brain or heart cell, comprising, e.g., contacting said cell with an agent effective to modulate OTB860, or the biological activity of a polypeptide encoded thereby (e.g., SEQ ID NO 39), or a mammalian homolog thereof, whereby said brain or heart cell is modulated. Modulation as used throughout includes, e.g., stimulating, increasing, agonizing, activating, amplifying, blocking, inhibiting, reducing, antagonizing, preventing, decreasing, diminishing, etc. Any activity or function of the brain or heart cell can be modulated, including, e.g., development, differentiation, signaling, excitability, etc.

The present invention also relates to methods of modulating development of cardiac or neuronal cells, comprising, e.g., administering an agent which is effective for modulating the expression of OTB860, or the biological activity of a polypeptide encoded thereby, whereby the development of said cardiac or neuronal cell is modulated. Development is meant to include any process in which a cell or tissue matures, including differentiation, organogenesis, cell proliferation, cell survival, expression and induction of functional molecules, cell movement and migration, apoptosis, modulation of gene expression, trabeculation, etc. Examples of heart and brain development that can be modulation are disclosed in Kramer et al., *Proc. Natl. Acad. Sci.*, 93:4833-4838, 1996. Any agent can be used to modulate development in any environment, e.g., in situ, in vivo, or in vitro.

OTB860 can be used to detect, modulate, etc., any of the cell types present in the heart or brain, including, but not limited to, heart cells comprising the coverings of the heart (e.g., pericardium, fibrous pericardium, serous pericardium containing the parietal layer and epicardium), heart wall comprising myocardium, cardiac muscle, and endocardium (endothelial), blood vessels, valves, and autorhythmic cardiac cells such as those in the SA and VA nodes, brain and other neuronal cells, such neurons, glia, microglia, ependymal cells,

-24-

oligodendrocytes, Schwann cells, and satellite cells.

Promoter sequences obtained from OTB860 can be utilized to selectively express heterologous genes in brain or heart cells. Methods of expressing a heterologous polynucleotide in brain or heart cells can comprise, e.g., expressing a nucleic acid construct in brain or heart cells, said construct comprising a promoter sequence operably linked to said heterologous polynucleotide, wherein said promoter sequence is obtained from OTB860, e.g., on genomic NT_022115.8. In addition to the cell lines mentioned below, the construct can be expressed in primary cells or in established cell lines.

The present invention also relates to methods of modulating development of cardiac or neuronal cells, comprising, e.g., administering an agent which is effective for modulating the expression of OTB60, or the biological activity of a polypeptide encoded thereby, whereby the development of said cardiac or neuronal cell is modulated.

The present invention also relates to a mammalian cell whose genome comprises a functional disruption of the human OTB860 gene within a polynucleotide sequence coding for amino acid residues 1-369 (SEQ ID NO 39) or 1546-1574 (SEQ ID NO 39). A non-human, transgenic mammal comprising such a cell can have a heart or neuronal tissue defect.

Antibodies can be produced, e.g., an antibody which is specific for: an epitope comprising amino acid 847, amino acid 867, or an epitope contained with amino acids 1-369 (SEQ ID NO 39), or amino acids 1546-1574 (SEQ ID NO 39).

TARPP

Human TARPP (thymocyte cyclic AMP regulated phosphoprotein, or, Br137A, B, C, D, and E) is represented by a family of alternative splice variants. Figs. 10-12 summarize the differences between the multiple forms. Br137E is an 847 amino acid polypeptide. Its nucleotide and amino acid sequences are shown in SEQ ID NOS 43 and 44. Br137B (SEQ ID NO 47 and 48) has a deletion of amino acids 267-300, Br137A (SEQ ID NO 45 and 46) has a deletion of amino acids 312-331, and Br137C (SEQ ID NO 49 and 50) has a deletion of both these domains. Br137D (SEQ ID NO 51 and 52) contains only the first 87 amino acids followed by a two-amino acid N-terminus which differs from the other forms. A partial clone, AL133109 (SEQ ID NO 55) as shown in Fig. 10, is missing the first 161 amino acids of Br137E, as well as having an amino acid difference at position 312 (SEQ ID NO 44).

Br137E contains a nuclear localization signal at about amino acids 107-124, an R3H domain (single-stranded nucleic acid binding domain) at about amino acids 147-224, and a proline rich region at about amino acids 476-682. These domains are also present in the A-C splice forms, but at different amino acid positions. Human TARPP has nucleic acid binding activity conferred by the corresponding binding domain indicating that it can bind nucleic acids, preferably single-stranded DNA or RNA. This binding activity can be assayed routinely, e.g., using gel electrophoresis band shift assays, e.g., as carried out in, e.g., U.S. Pat. No. 6,333,407 and 5,789,538, ELISA-based assays (e.g., Mercury™ TransFactor Kit from Clontech), and other assays which detect DNA-protein interactions.

The Br137 family represent the human homologs of murine TARPP (thymocyte ARPP) (NM_033264; SEQ ID NO 53; "Mouse" in Fig. 12). Br137E has about 83% amino acid identity and 87% homology with it (calculated using the BLAST algorithm). See, Fig. 12 (NM_033264 is murine TARPP). In addition to amino acid sequence differences between the two proteins, human TARPP has an insertion at about amino acid positions 549-572 of SEQ ID NO 44 which is not present in the mouse protein. See, Fig. 12.

Originally, a 21 kDa polypeptide was isolated from rat basal ganglia based on its phosphorylation by cAMP-dependent protein kinase (PKA). Williams et al., *J. Neurosci.*, 9:3631-3637, 1989. It was named ARPP-21 (cAMP-regulated phosphoprotein). Activation of dopamine receptors resulted in an increase in the phosphorylation of ARPP-21. Caporaso et al., *Neuropharm.*, 39:1637-1644, 2000. Human ARPP-21 (Br137D) contains 89 amino acids (NM_016300; SEQ ID NO 52).

A high molecular weight polypeptide of ARPP-21 was subsequently identified in T-cells and named TARPP. Kisielow et al., *Eur. J. Immunol.*, 31:1141-1149, 2001. This polypeptide contains ARPP-21 sequence at its 5' end, but a novel 3' end coding for more than 700 additional amino acids (for a total of 807 amino acids). Murine TARPP appears to be involved in the regulation of thymocyte maturation and TCR rearrangement. Expression of TARPP is down-regulated after the TCR signals delivered. It is highly expressed in immature thymocytes and is associated with the commitment to the T-cell lineage, making it highly selective marker for T-cell commitment. See, Kisielow, *ibid.* After commitment to the T-cell lineage during positive selection, its expression is turned off.

There appear to be several members of the human TARPP family. KIAA0029 is a

-26-

hypothetical protein that shares about 45% amino acid sequence identity and 59% homology with Br137E. KIAA1002, a second hypothetical protein, has about 42% amino acid identity and 54% homology to it.

Human TARPP is highly expressed in brain, pituitary, muscle, and thymus. It is expressed at lower levels in adrenal gland, bone marrow, heart, small intestine, kidney, liver, ovary, prostate, stomach, testis, and thyroid. There was virtually no detectable expression in colon, lung, lymph node, peripheral lymphocytes, mammary gland, pancreas, and uterus.

As indicated by its expression pattern, human TARPP is involved the maturation of T-cells, especially in the rearrangement of the TCR. For this reason, it can be used to modulate T-cells, e.g., in allergy, autoimmune disease (e.g., rheumatoid arthritis and multiple sclerosis), and graft-host disease. It can also be used a marker to determine the index of mature versus immature T-cells, where human TARPP is marker of immature T-cells. Additionally, human TARPP is phosphorylated upon dopamine receptor activation, indicating an involvement in dopamine pathways. Consequently, it is target for diseases that involve dopamine, including, e.g., schizophrenia, substance abuse and addiction, anxiety, Parkinson's disease, and other dopaminergic diseases and conditions.

Human TARPP is localized to chromosomal band 3p21.33. There are several disorders genetically mapped to this region, including, e.g., retinal vasculopathy with cerebral leukodystrophy (OMIM 192315), deafness, neurosensory, autosomal recessive 6 (OMIM 600971), and lung cancer. Nucleic acids of the present invention can be used as linkage markers, diagnostic targets, therapeutic targets, for any of the mentioned disorders, as well as any disorders or genes mapping in proximity to it.

Diseases or disorders which can be treated in accordance with the present invention include, but are not limited to autoimmune disease, such as multiple sclerosis and rheumatoid arthritis, and allergy

The gene can be disrupted in a specific region, e.g., in the sequence coding for amino acids 1-161 of a human TARPP. Cells and/or animals can also have targeted deletions, e.g., deletion of a coding sequence for amino acids 267-300 and/or 312-331 of a human TARPP. One or more the different splice forms, Br137A-E can also be knocked-out or disrupted, e.g., to dissect out the individual activities.

-27-

The present invention relates to methods of modulating T-cells, comprising, contacting T-cells with an agent which is effective for regulating a human TARPP gene expressed in said cells, or for modulating the biological activity of a polypeptide encoded thereby.

5 Included also in the present invention are engineered cells, e.g., a human cell whose genome comprises a functional disruption of human TARPP in the region comprising the coding sequence for amino acids 1-161 of a human TARPP of SEQ ID NO 44, or a human cell whose genome comprises a deletion of a coding sequence for amino acids 267-300 and/or 312-331 of a human TARPP of SEQ ID NO 44.

10 Antibodies can be produced, e.g., an antibody which is specific-for a human TARPP, said antibody which is specific for an epitope present in amino acid sequences 1-161, 88-161, 267-300, 312-331, or a polypeptide comprising amino acid 312, of a human TARPP of SEQ ID NO 44.

15 LAT-1

Liver-associated transmembrane protein-1 ("LAT-1" or TMD008) codes for a polypeptide comprising 276 amino acids. Its expression is highly restricted to the liver, i.e., it is predominantly expressed in the liver. The nucleotide and amino acid sequences of it are shown in SEQ ID NOS 58 and 59. It contains transmembrane domains at about amino acid
20 positions 24-46, 59-81, 101-123, 144-166, 203-225, and 237-259. It is homologous to the olfactory class of GPCR receptors. LAT-1 is also known as XM_060456 and AX242289.

The gene for LAT-1 maps to chromosomal band 1q22. Several different disorders map to this location, including, e.g., porphyria variegata, progression of lymphoma, Zellweger syndrome, Charcot-Marie-Tooth neuropathy-1B, congenital hypomyelination,
25 nemaline myopathy, and CD3 zeta chain deficiency, medullary thyroid carcinoma, susceptibility to Vivax malaria, schizophrenia susceptibility locus, autosomal dominant deafness, susceptibility to Lupus nephritis, familial hemiplegic migraine, apolipoprotein A-II deficiency, and familial hyperlipidemia. Nucleic acids of the present invention can be used as linkage markers, diagnostic targets, therapeutic targets, for any of the mentioned disorders,
30 as well as any disorders or genes mapping in proximity to it.

LAT-1 can be used as a diagnostic and prognostic marker for liver function and

-28-

disease, including any of the liver diseases already mentioned. For instance, blood serum levels of LAT-1 (as well as other bodily fluids) can be used as an indicator of liver disease, especially those diseases characterized by necrotic and degenerative lesions, such as hepatitis, toxicity, and cirrhosis. Any condition which results in degeneration of the liver can result in the appearance of higher than normal amounts of blood serum LAT-1. LAT-1 can be used alone, or in combination with other molecular markers for liver function, such as bilirubin, serum aminotransferases (e.g., AST and ALT), alkaline phosphatase, gamma-glutamyltranspeptidase (GGT), albumin, globulin, and blood ammonia.

Because of the selectivity of LAT-1 for the liver, it is a useful target for both histological and therapeutic applications. Antibodies and other LAT-1 binding partners can be used to selectively target agents to liver tissue for any purpose, included, but not limited to, imaging, therapeutic, diagnostic, drug delivery, gene therapy, etc. For example, LAT-1 binding partners, such as antibodies, can be used to treat liver carcinoma, in analogy to how c-erbB-2 antibodies are used to breast cancer, to detect metastatic liver cells, etc. Useful antibodies or other binding partners include those that are specific for parts of LAT-1 which are exposed extracellularly, e.g., amino acids 1-23, 82-100, 167-202, etc.

Imaging of specific organs can be facilitated using agents, such as LAT-1, that can be used to selectively target contrast agents to a specific site in the body. Various imaging techniques have been used in this context, including, e.g., X-ray, CT, CAT, MRI, ultrasound, PET, SPECT, and scintigraphic. A reporter agent can be conjugated or associated routinely with a LAT-1 binding partner. Ultrasound contrast agents combined with binding partners, such as antibodies, are described in, e.g., U.S. Pat. Nos. 6,264,917, 6,254,852, 6,245,318, and 6,139,819. MRI contrast agents, such as metal chelators, radionucleotides, paramagnetic ions, etc., combined with selective targeting agents are also described in the literature, e.g., in U.S. Pat. Nos. 6,280,706 and 6,221,334. The methods described therein can be used generally to associate a LAT-1 binding partner with an agent for any desired purpose.

LAT-1 binding partners can also be used as to specifically deliver therapeutic agents to the liver. For example, hypercholesterolemia and other metabolic diseases can be treated by gene therapy, using the LAT-1 to specifically deliver the LDL receptor to the liver. The gene can be conjugated to a LAT-1 binding partner (directly or through a polymer, etc.), in liposomes comprising cell surface. Additionally, cytotoxic, cytostatic, and other therapeutic

agents can be delivered to the liver via LAT-1 to treat and/or prevent any of the above-mentioned conditions associated with liver disease, e.g., carcinoma.

The liver is the largest and most metabolically complex organ in the body. Its functions include, e.g., storage of iron, production of bile to facilitate digestion, detoxifications of various exogenous chemicals, including alcohol and many drugs, energy stockpiling (carbohydrates and fat), production of clotting factors, and manufacture of blood. There are a number of diseases which affect the liver, including, Alagille syndrome, alcoholic liver disease, alpha-1-antitrypsin deficiency, autoimmune hepatitis, Budd-Chiari syndrome, biliary atresia, Byler disease, liver cancer, Caroli disease, cirrhosis, Crigler-Najjar syndrome, Dubin-Johnson syndrome, fatty liver, galactosemia, Gilbert syndrome, glycogen storage disease, hemangioma, hemochromatosis, hepatitis A-G, porphyria cutanea tarda, primary biliary cirrhosis, protoporphyria, erythrohepatic, Rotor syndrome, sclerosing cholangitis, and Wilson disease. Liver disease is of grave concern around the world.

The liver is divided into many small units, known as lobules. The lobule is the structural unit of the liver. Each lobule is comprised of radial plates of liver cells, called hepatocytes, and is surrounded by a connective sheath. A central vein ("CV") is located in the middle, and there are portal triads at the vertices. Each triad comprises a branch of the hepatic artery (supplying arterial blood to the lobule), a branch of the hepatic portal vein (carrying nutrient-rich blood from the digestive viscera), and a bile duct. The blood from the artery and portal vein flow into leaky capillaries, the liver sinusoids, which are located between the hepatic plates of the lobule.

The acinus is the functional unit of the liver. While the boundaries of the lobule are well visible, those of the acinus are unrecognizable under the microscope. Arising like a berry, a grape (latin "acinus") on the vine around the portal triad, the liver acinus is formed of a mass of liver cells and sinusoids which drain toward two adjacent central veins. The principal metabolic functions of the liver are performed by hepatocytes. These functions include, e.g., formation and excretion of bile, regulation of carbohydrate homeostasis, lipid synthesis and secretion of plasma lipoproteins, regulation of cholesterol metabolism, formation of urea, serum albumin, clotting factors, enzymes, and numerous other proteins; and metabolism or detoxification of drugs and other foreign substances. Hepatocytes in different regions of the acinus perform different functions, e.g., gluconeogenesis is primarily

a function of the zone of cells closest to the triad, whereas glycolysis mainly occurs in the farthest zone from it.

A promoter obtained from the LAT-1 can be used, e.g., in gene therapy to obtain tissue-specific expression of a heterologous gene (e.g., coding for a therapeutic product or cytotoxin). A promoter sequence is found at about nucleotide positions 1164-1212 of SEQ ID NO 58 and can be used (e.g., 1164 to the first ATG codon) to drive liver-specific expression of a heterologous sequence. 5' and 3' sequences (including, UTRs and introns) can be used to modulate or regulate stability, transcription, and translation of nucleic acids, including the sequence to which is attached in nature, as well as heterologous nucleic acids. A polyadenylation site is found at about nucleotide positions 4265-4275 of SEQ ID NO 58. The upstream 3'UTR can be used as described above. Useful polypeptides include polypeptides exposed extracellularly, e.g., amino acids 1-23, 82-100, 167-202, of SEQ ID NO 59, etc.

The present invention also relates to methods of detecting human liver tissue in a sample, e.g., comprising tissue, cells, or other cellular materials or debris, comprising one or more of the following steps, e.g., contacting said sample with a binding partner specific for human LAT-1 under conditions effective for said binding partner to bind specifically to human LAT-1, and detecting specific binding between said binding partner and said human LAT-1, whereby specific binding indicates that liver tissue is present in said sample.

The sample can be contacted with the binding partner in any manner which is effective to give the binding partner access to the material present in the tissue sample. How contact is achieved can depend on the format of the detection assay. For instance, if a ELISA assay is used, and the binding partner is an antibody on a solid phase in a well, then placing an aqueous sample in the well would achieve contact between partner and sample. Any type of sample can be used, including, e.g., blood (whole blood, fractionated blood, serum, etc.), stool, urine, cerebral spinal fluid, tissue biopsy, etc.

The binding partner, such as a monoclonal or polyclonal antibody, is specific for LAT-1, and is contacted with the sample under conditions effective for said binding partner to bind specifically to human LAT-1, if human LAT-1 is present in the sample. Specific binding, as previously discussed for polynucleotides, indicates that the binding partner binds or attaches to its target polypeptide without significant binding to other polypeptides ("non-

-31-

specific binding'") in the sample. This concept is well known in the art. The detection of specific binding can be accomplished by any of the aforementioned assays.

The present invention also relates to polypeptide detection methods for assessing liver function, e.g., methods of assessing liver function, comprising, detecting LAT-1 polypeptide, or fragments thereof, in a body fluid, whereby the level of LAT-1 polypeptide in said fluid is a measure of liver function. Liver function tests are usually performed to determine whether the liver is functioning normally as a way of diagnosing liver disease. Various tests are commonly used, including, e.g., alkaline phosphatase, alanine transferase, aspartate transferase, bilirubin, gamma-glutamyl transpeptidase, lactic dehydrogenase, 5'-nucleotidase, albumin, alpha-fetoprotein, mitochondrial antibodies, and prothrombin time. See, e.g., *Harrison's Principles of Internal Medicine*, Volume 2, Pages 1308-1317, 12th Edition, 1991. Detection of LAT-1 provides an additional assessment tool, especially in diseases such as hepatitis, carcinoma, liver toxicity, cirrhosis, and other liver conditions, e.g., where cellular debris, etc., is released systemically. As with the other tests, elevated levels of LAT-1 in blood, or other fluids, can indicate impaired liver function. Values can be determined routinely, as they are for other liver function markers.

Nucleic acids

A mammalian polynucleotide, or fragment thereof, of the present invention is a polynucleotide having a nucleotide sequence obtainable from a natural source. When the species name is used, e.g., a human, it indicates that the polynucleotide or polypeptide is obtainable from a natural source. It therefore includes naturally-occurring normal, naturally-occurring mutant, and naturally-occurring polymorphic alleles (e.g., SNPs), differentially-spliced transcripts, splice-variants, etc. By the term "naturally-occurring," it is meant that the polynucleotide is obtainable from a natural source, e.g., animal tissue and cells, body fluids, tissue culture cells, forensic samples. Natural sources include, e.g., living cells obtained from tissues and whole organisms, tumors, cultured cell lines, including primary and immortalized cell lines. Naturally-occurring mutations can include deletions (e.g., a truncated amino- or carboxy-terminus), substitutions, inversions, or additions of nucleotide sequence. These genes can be detected and isolated by polynucleotide hybridization according to methods which one skilled in the art would know, e.g., as discussed below.

-32-

A polynucleotide according to the present invention can be obtained from a variety of different sources. It can be obtained from DNA or RNA, such as polyadenylated mRNA or total RNA, e.g., isolated from tissues, cells, or whole organism. The polynucleotide can be obtained directly from DNA or RNA, from a cDNA library, from a genomic library, etc. The polynucleotide can be obtained from a cell or tissue (e.g., from an embryonic or adult tissues) at a particular stage of development, having a desired genotype, phenotype, disease status, etc. A polynucleotide which "codes without interruption" refers to a polynucleotide having a continuous open reading frame ("ORF") as compared to an ORF which is interrupted by introns or other noncoding sequences.

Polynucleotides and polypeptides (including any part of a differentially regulated cancer gene) can be excluded as compositions from the present invention if, e.g., listed in a publicly available databases on the day this application was filed and/or disclosed in a patent application having an earlier filing or priority date than this application and/or conceived and/or reduced to practice earlier than a polynucleotide in this application.

As described herein, the phrase "an isolated polynucleotide which is SEQ ID NO," or "an isolated polynucleotide which is selected from SEQ ID NO," refers to an isolated nucleic acid molecule from which the recited sequence was derived (e.g., a cDNA derived from mRNA; cDNA derived from genomic DNA). Because of sequencing errors, typographical errors, etc., the actual naturally-occurring sequence may differ from a SEQ ID listed herein.

Thus, the phrase indicates the specific molecule from which the sequence was derived, rather than a molecule having that exact recited nucleotide sequence, analogously to how a culture depository number refers to a specific cloned fragment in a cryotube.

As explained in more detail below, a polynucleotide sequence of the invention can contain the complete sequence as shown in the corresponding SEQ ID, degenerate sequences thereof, anti-sense, muteins thereof, genes comprising said sequences, full-length cDNAs comprising said sequences, complete genomic sequences, fragments thereof, homologs, primers, nucleic acid molecules which hybridize thereto, derivatives thereof, etc.

Genomic

The present invention also relates genomic DNA from which the polynucleotides of the present invention can be derived. A genomic DNA coding for a human, mouse, or other

-33-

mammalian polynucleotide, can be obtained routinely, for example, by screening a genomic library (e.g., a YAC library) with a polynucleotide of the present invention, or by searching nucleotide databases, such as GenBank and EMBL, for matches. Promoter and other regulatory regions (including both 5' and 3' regions, as well introns) can be identified upstream or downstream of coding and expressed RNAs, and assayed routinely for activity, e.g., by joining to a reporter gene (e.g., CAT, GFP, alkaline phosphatase, luciferase, galactosidase). A promoter obtained from a gene can be used, e.g., in gene therapy to obtain tissue-specific expression of a heterologous gene (e.g., coding for a therapeutic product or cytotoxin). 5' and 3' sequences (including, UTRs and introns) can be used to modulate or regulate stability, transcription, and translation of nucleic acids, including the sequence to which is attached in nature, as well as heterologous nucleic acids.

Constructs

A polynucleotide of the present invention can comprise additional polynucleotide sequences, e.g., sequences to enhance expression, detection, uptake, cataloging, tagging, etc. A polynucleotide can include only coding sequence; a coding sequence and additional non-naturally occurring or heterologous coding sequence (e.g., sequences coding for leader, signal, secretory, targeting, enzymatic, fluorescent, antibiotic resistance, and other functional or diagnostic peptides); coding sequences and non-coding sequences, e.g., untranslated sequences at either a 5' or 3' end, or dispersed in the coding sequence, e.g., introns.

A polynucleotide according to the present invention also can comprise an expression control sequence operably linked to a polynucleotide as described above. The phrase "expression control sequence" means a polynucleotide sequence that regulates expression of a polypeptide coded for by a polynucleotide to which it is functionally ("operably") linked. Expression can be regulated at the level of the mRNA or polypeptide. Thus, the expression control sequence includes mRNA-related elements and protein-related elements. Such elements include promoters, enhancers (viral or cellular), ribosome binding sequences, transcriptional terminators, etc. An expression control sequence is operably linked to a nucleotide coding sequence when the expression control sequence is positioned in such a manner to effect or achieve expression of the coding sequence. For example, when a promoter is operably linked 5' to a coding sequence, expression of the coding sequence is

-34-

driven by the promoter. Expression control sequences can include an initiation codon and additional nucleotides to place a partial nucleotide sequence of the present invention in-frame in order to produce a polypeptide (e.g., pET vectors from Promega have been designed to permit a molecule to be inserted into all three reading frames to identify the one that results in polypeptide expression). Expression control sequences can be heterologous or endogenous to the normal gene.

A polynucleotide of the present invention can also comprise nucleic acid vector sequences, e.g., for cloning, expression, amplification, selection, etc. Any effective vector can be used. A vector is, e.g., a polynucleotide molecule which can replicate autonomously in a host cell, e.g., containing an origin of replication. Vectors can be useful to perform manipulations, to propagate, and/or obtain large quantities of the recombinant molecule in a desired host. A skilled worker can select a vector depending on the purpose desired, e.g., to propagate the recombinant molecule in bacteria, yeast, insect, or mammalian cells. The following vectors are provided by way of example. Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBS, pD10, Phagescript, phiX174, pBK Phagemid, pNH8A, pNH16a, pNH18Z, pNH46A (Stratagene); Bluescript KS+II (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR54 0, pRIT5 (Pharmacia). Eukaryotic: PWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene), pSVK3, PBPV, PMSG, pSVL (Pharmacia), pCR2.1/TOPO, pCRII/TOPO, pCR4/TOPO, pTrcHisB, pCMV6-XL4, etc. However, any other vector, e.g., plasmids, viruses, or parts thereof, may be used as long as they are replicable and viable in the desired host. The vector can also comprise sequences which enable it to replicate in the host whose genome is to be modified.

Hybridization

Polynucleotide hybridization, as discussed in more detail below, is useful in a variety of applications, including, in gene detection methods, for identifying mutations, for making mutations, to identify homologs in the same and different species, to identify related members of the same gene family, in diagnostic and prognostic assays, in therapeutic applications (e.g., where an antisense polynucleotide is used to inhibit expression), etc.

The ability of two single-stranded polynucleotide preparations to hybridize together is a measure of their nucleotide sequence complementarity, e.g., base-pairing between

-35-

nucleotides, such as A-T, G-C, etc. The invention thus also relates to polynucleotides, and their complements, which hybridize to a polynucleotide comprising a nucleotide sequence as set forth in the sequences disclosed herein, and genomic sequences thereof. A nucleotide sequence hybridizing to the latter sequence will have a complementary polynucleotide strand, or act as a template for one in the presence of a polymerase (i.e., an appropriate polynucleotide synthesizing enzyme). The present invention includes both strands of polynucleotide, e.g., a sense strand and an anti-sense strand.

Hybridization conditions can be chosen to select polynucleotides which have a desired amount of nucleotide complementarity with the nucleotide sequences set forth in SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58 and genomic sequences thereof. A polynucleotide capable of hybridizing to such sequence, preferably, possesses, e.g., about 70%, 75%, 80%, 85%, 87%, 90%, 92%, 95%, 97%, 99%, or 100% complementarity, between the sequences. The present invention particularly relates to polynucleotide sequences which hybridize to the nucleotide sequences set forth in the sequence disclosure herein or genomic sequences thereof, under low or high stringency conditions. These conditions can be used, e.g., to select corresponding homologs in non-human species.

Polynucleotides which hybridize to polynucleotides of the present invention can be selected in various ways. Filter-type blots (i.e., matrices containing polynucleotide, such as nitrocellulose), glass chips, and other matrices and substrates comprising polynucleotides (short or long) of interest, can be incubated in a prehybridization solution (e.g., 6X SSC, 0.5% SDS, 100 µg/ml denatured salmon sperm DNA, 5X Denhardt's solution, and 50% formamide), at 22-68°C, overnight, and then hybridized with a detectable polynucleotide probe under conditions appropriate to achieve the desired stringency. In general, when high homology or sequence identity is desired, a high temperature can be used (e.g., 65 °C). As the homology drops, lower washing temperatures are used. For salt concentrations, the lower the salt concentration, the higher the stringency. The length of the probe is another consideration. Very short probes (e.g., less than 100 base pairs) are washed at lower temperatures, even if the homology is high. With short probes, formamide can be omitted. See, e.g., *Current Protocols in Molecular Biology*, Chapter 6, Screening of Recombinant Libraries; Sambrook et al., *Molecular Cloning*, 1989, Chapter 9.

-36-

For instance, high stringency conditions can be achieved by incubating the blot overnight (e.g., at least 12 hours) with a polynucleotide probe in a hybridization solution containing, e.g., about 5X SSC, 0.5% SDS, 100 µg/ml denatured salmon sperm DNA and 50% formamide, at 42°C, or hybridizing at 42°C in 5X SSPE, 0.5% SDS, and 50%
5 formamide, 100 µg/ml denatured salmon sperm DNA, and washing at 65°C in 0.1X SSC and 0.1% SDS.

Blots can be washed at high stringency conditions that allow, e.g., for less than 5% bp mismatch (e.g., wash twice in 0.1% SSC and 0.1% SDS for 30 min at 65°C), i.e., selecting sequences having 95% or greater sequence identity.

10 Other non-limiting examples of high stringency conditions includes a final wash at 65°C in aqueous buffer containing 30 mM NaCl and 0.5% SDS. Another example of high stringent conditions is hybridization in 7% SDS, 0.5 M NaPO₄, pH 7, 1 mM EDTA at 50°C, e.g., overnight, followed by one or more washes with a 1% SDS solution at 42°C.

Whereas high stringency washes can allow for, e.g., less than 10%, less than 5%
15 mismatch, etc., reduced or low stringency conditions can permit up to 20% nucleotide mismatch. Hybridization at low stringency can be accomplished as above, but using lower formamide conditions, lower temperatures and/or lower salt concentrations, as well as longer periods of incubation time.

Hybridization can also be based on a calculation of melting temperature (T_m) of the
20 hybrid formed between the probe and its target, as described in Sambrook et al.. Generally, the temperature T_m at which a short oligonucleotide (containing 18 nucleotides or fewer) will melt from its target sequence is given by the following equation: $T_m = (\text{number of A's and T's}) \times 2^\circ\text{C} + (\text{number of C's and G's}) \times 4^\circ\text{C}$. For longer molecules, $T_m = 81.5 + 16.6 \log_{10}[\text{Na}^+] + 0.41(\%GC) - 600/N$ where $[\text{Na}^+]$ is the molar concentration of sodium ions,
25 %GC is the percentage of GC base pairs in the probe, and N is the length. Hybridization can be carried out at several degrees below this temperature to ensure that the probe and target can hybridize. Mismatches can be allowed for by lowering the temperature even further.

Stringent conditions can be selected to isolate sequences, and their complements, which have, e.g., at least about 90%, 95%, or 97%, nucleotide complementarity between the
30 and a target polynucleotide.

-37-

Other homologs of polynucleotides of the present invention can be obtained from mammalian and non-mammalian sources according to various methods. For example, hybridization with a polynucleotide can be employed to select homologs, e.g., as described in Sambrook et al., *Molecular Cloning*, Chapter 11, 1989. Such homologs can have varying amounts of nucleotide and amino acid sequence identity and similarity to such polynucleotides of the present invention. Mammalian organisms include, e.g., mice, rats, monkeys, pigs, cows, etc. Non-mammalian organisms include, e.g., vertebrates, invertebrates, zebra fish, chicken, *Drosophila*, *C. elegans*, *Xenopus*, yeast such as *S. pombe*, *S. cerevisiae*, roundworms, prokaryotes, plants, *Arabidopsis*, *artemia*, viruses, etc. The degree of nucleotide sequence identity between human and mouse can be about, e.g. 70% or more, 85% or more for open reading frames, etc.

Alignment

Alignments can be accomplished by using any effective algorithm. For pairwise alignments of DNA sequences, the methods described by Wilbur-Lipman (e.g., Wilbur and Lipman, *Proc. Natl. Acad. Sci.*, 80:726-730, 1983) or Martinez/Needleman-Wunsch (e.g., Martinez, *Nucleic Acid Res.*, 11:4629-4634, 1983) can be used. For instance, if the Martinez/Needleman-Wunsch DNA alignment is applied, the minimum match can be set at 9, gap penalty at 1.10, and gap length penalty at 0.33. The results can be calculated as a similarity index, equal to the sum of the matching residues divided by the sum of all residues and gap characters, and then multiplied by 100 to express as a percent. Similarity index for related genes at the nucleotide level in accordance with the present invention can be greater than 70%, 80%, 85%, 90%, 95%, 99%, or more. Pairs of protein sequences can be aligned by the Lipman-Pearson method (e.g., Lipman and Pearson, *Science*, 227:1435-1441, 1985) with k-tuple set at 2, gap penalty set at 4, and gap length penalty set at 12. Results can be expressed as percent similarity index, where related genes at the amino acid level in accordance with the present invention can be greater than 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99%, or more. Various commercial and free sources of alignment programs are available, e.g., MegAlign by DNA Star, BLAST (National Center for Biotechnology Information), BCM (Baylor College of Medicine) Launcher, etc. BLAST can be used to calculate amino acid sequence identity, amino acid sequence homology, and nucleotide

-38-

sequence identity. These calculations can be made along the entire length of each of the target sequences which are to be compared.

After two sequences have been aligned, a "percent sequence identity" can be determined. For these purposes, it is convenient to refer to a Reference Sequence and a

- 5 Compared Sequence, where the Compared Sequence is *compared* to the Reference Sequence. Percent sequence identity can be determined according to the following formula: Percent Identity = $100 [1 - (C/R)]$, wherein C is the number of differences between the Reference Sequence and the Compared Sequence over the length of alignment between the Reference Sequence and the Compared Sequence where (i) each base or amino acid in the Reference
- 10 Sequence that does not have a corresponding aligned base or amino acid in the Compared Sequence, (ii) each gap in the Reference Sequence, (iii) each aligned base or amino acid in the Reference Sequence that is different from an aligned base or amino acid in the Compared Sequence, constitutes a difference; and R is the number of bases or amino acids in the Reference Sequence over the length of the alignment with the Compared Sequence with any
- 15 gap created in the Reference Sequence also being counted as a base or amino acid. When it is stated that a polynucleotide sequence has a certain percentage having 95% or more sequence identity along the entire length of the polynucleotide sequence set forth in SEQ ID NO 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, or 58, it means that when the polynucleotide is shorter than the mentioned SEQ ID NOS, the missing bases are counted for the purposes of the calculation.
- 20 Percent sequence identity can also be determined by other conventional methods, e.g., as described in Altschul et al., *Bull. Math. Bio.* 48: 603-616, 1986 and Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915-10919, 1992.

Specific polynucleotide probes

- 25 A polynucleotide of the present invention can comprise any continuous nucleotide sequence of SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58, sequences which share sequence identity thereto, or complements thereof. The term "probe" refers to any substance that can be used to detect, identify, isolate, etc., another substance. A polynucleotide probe is comprised of nucleic acid can be used to detect, identify, etc., other
- 30 nucleic acids, such as DNA and RNA.

These polynucleotides can be of any desired size that is effective to achieve the specificity desired. For example, a probe can be from about 7 or 8 nucleotides to several thousand nucleotides, depending upon its use and purpose. For instance, a probe used as a primer PCR can be shorter than a probe used in an ordered array of polynucleotide probes.

5 Probe sizes vary, and the invention is not limited in any way by their size, e.g., probes can be from about 7-2000 nucleotides, 7-1000, 8-700, 8-600, 8-500, 8-400, 8-300, 8-150, 8-100, 8-75, 7-50, 10-25, 14-16, at least about 8, at least about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, or more, etc. The polynucleotides can have non-naturally-occurring nucleotides, e.g., inosine, AZT, 3TC, etc. The polynucleotides can have 100% sequence
10 identity or complementarity to a sequence of SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58, or it can have mismatches or nucleotide substitutions, e.g., 1, 2, 3, 4, or 5 substitutions. The probes can be single-stranded or double-stranded.

In accordance with the present invention, a polynucleotide can be present in a kit, where the kit includes, e.g., one or more polynucleotides, a desired buffer (e.g., phosphate,
15 tris, etc.), detection compositions, RNA or cDNA from different tissues to be used as controls, libraries, etc. The polynucleotide can be labeled or unlabeled, with radioactive or non-radioactive labels as known in the art. Kits can comprise one or more pairs of polynucleotides for amplifying nucleic acids specific for polynucleotides, e.g., comprising a forward and reverse primer effective in PCR. These include both sense and anti-sense
20 orientations. For instance, in PCR-based methods (such as RT-PCR), a pair of primers are typically used, one having a sense sequence and the other having an antisense sequence.

Another aspect of the present invention is a nucleotide sequence that is specific to, or for, a selective polynucleotide. The phrases "specific for" or "specific to" a polynucleotide have a functional meaning that the polynucleotide can be used to identify the presence of one
25 or more target genes in a sample and distinguish them from non-target genes. It is specific in the sense that it can be used to detect polynucleotides above background noise ("non-specific binding"). A specific sequence is a defined order of nucleotides (or amino acid sequences, if it is a polypeptide sequence) which occurs in the polynucleotide, e.g., in the nucleotide sequences of SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58, and which is
30 characteristic of that target sequence, and substantially no non-target sequences. A probe or mixture of probes can comprise a sequence or sequences that are specific to a plurality of

target sequences, e.g., where the sequence is a consensus sequence, a functional domain, etc., e.g., capable of recognizing a family of related genes. Such sequences can be used as probes in any of the methods described herein or incorporated by reference. Both sense and antisense nucleotide sequences are included. A specific polynucleotide according to the present invention can be determined routinely.

A polynucleotide comprising a specific sequence can be used as a hybridization probe to identify the presence of, e.g., human or mouse polynucleotide, in a sample comprising a mixture of polynucleotides, e.g., on a Northern blot. Hybridization can be performed under high stringent conditions (see, above) to select polynucleotides (and their complements which can contain the coding sequence) having at least 90%, 95%, 99%, etc., identity (i.e., complementarity) to the probe, but less stringent conditions can also be used. A specific polynucleotide sequence can also be fused in-frame, at either its 5' or 3' end, to various nucleotide sequences as mentioned throughout the patent, including coding sequences for enzymes, detectable markers, GFP, etc, expression control sequences, etc.

A polynucleotide probe, especially one that is specific to a polynucleotide of the present invention, can be used in gene detection and hybridization methods as already described. In one embodiment, a specific polynucleotide probe can be used to detect whether a particular tissue or cell-type is present in a target sample, e.g., with OTB182, OTB860, or LAT-1. To carry out such a method, a selective polynucleotide can be chosen which is characteristic of the desired target tissue. Such polynucleotide is preferably chosen so that it is expressed or displayed in the target tissue, but not in other tissues which are present in the sample. For instance, if detection of brain tissue is desired, it may not matter whether the selective polynucleotide is expressed in other tissues, as long as it is not expressed in cells normally present in blood, e.g., peripheral blood mononuclear cells. Starting from the selective polynucleotide, a specific polynucleotide probe can be designed which hybridizes (if hybridization is the basis of the assay) under the hybridization conditions to the selective polynucleotide, whereby the presence of the selective polynucleotide can be determined.

Probes which are specific for polynucleotides of the present invention can also be prepared using involve transcription-based systems, e.g., incorporating an RNA polymerase

promoter into a selective polynucleotide of the present invention, and then transcribing anti-sense RNA using the polynucleotide as a template. See, e.g., U.S. Pat. No. 5,545,522.

Polynucleotide composition

5 A polynucleotide according to the present invention can comprise, e.g., DNA, RNA, synthetic polynucleotide, peptide polynucleotide, modified nucleotides, dsDNA, ssDNA, ssRNA, dsRNA, and mixtures thereof. A polynucleotide can be single- or double-stranded, triplex, DNA:RNA, duplexes, comprise hairpins, and other secondary structures, etc. Nucleotides comprising a polynucleotide can be joined via various known linkages, e.g.,
10 ester, sulfamate, sulfamide, phosphorothioate, phosphoramidate, methylphosphonate, carbamate, etc., depending on the desired purpose, e.g., resistance to nucleases, such as RNase H, improved in vivo stability, etc. See, e.g., U.S. Pat. No. 5,378,825. Any desired nucleotide or nucleotide analog can be incorporated, e.g., 6-mercaptoguanine, 8-oxo-guanine, etc.

15 Various modifications can be made to the polynucleotides, such as attaching detectable markers (avidin, biotin, radioactive elements, fluorescent tags and dyes, energy transfer labels, energy-emitting labels, binding partners, etc.) or moieties which improve hybridization, detection, and/or stability. The polynucleotides can also be attached to solid supports, e.g., nitrocellulose, magnetic or paramagnetic microspheres (e.g., as described in
20 U.S. Pat. No. 5,411,863; U.S. Pat. No. 5,543,289; for instance, comprising ferromagnetic, supermagnetic, paramagnetic, superparamagnetic, iron oxide and polysaccharide), nylon, agarose, diazotized cellulose, latex solid microspheres, polyacrylamides, etc., according to a desired method. See, e.g., U.S. Pat. Nos. 5,470,967, 5,476,925, and 5,478,893.

Polynucleotide according to the present invention can be labeled according to any
25 desired method. The polynucleotide can be labeled using radioactive tracers such as ^{32}P , ^{35}S , ^3H , or ^{14}C , to mention some commonly used tracers. The radioactive labeling can be carried out according to any method, such as, for example, terminal labeling at the 3' or 5' end using a radiolabeled nucleotide, polynucleotide kinase (with or without dephosphorylation with a phosphatase) or a ligase (depending on the end to be labeled). A non-radioactive labeling can
30 also be used, combining a polynucleotide of the present invention with residues having immunological properties (antigens, haptens), a specific affinity for certain reagents

(ligands), properties enabling detectable enzyme reactions to be completed (enzymes or coenzymes, enzyme substrates, or other substances involved in an enzymatic reaction), or characteristic physical properties, such as fluorescence or the emission or absorption of light at a desired wavelength, etc.

5

Nucleic acid detection methods

Another aspect of the present invention relates to methods and processes for detecting polynucleotides. Detection methods have a variety of applications, including for diagnostic, prognostic, forensic, and research applications. To accomplish gene detection, a
10 polynucleotide in accordance with the present invention can be used as a "probe." The term "probe" or "polynucleotide probe" has its customary meaning in the art, e.g., a polynucleotide which is effective to identify (e.g., by hybridization), when used in an appropriate process, the presence of a target polynucleotide to which it is designed. Identification can involve simply determining presence or absence, or it can be quantitative, e.g., in assessing amounts
15 of a gene or gene transcript present in a sample. Probes can be useful in a variety of ways, such as for diagnostic purposes, to identify homologs, and to detect, quantitate, or isolate a polynucleotide of the present invention in a test sample.

Assays can be utilized which permit quantification and/or presence/absence detection of a target nucleic acid in a sample. Assays can be performed at the single-cell level, or in a
20 sample comprising many cells, where the assay is "averaging" expression over the entire collection of cells and tissue present in the sample. Any suitable assay format can be used, including, but not limited to, e.g., Southern blot analysis, Northern blot analysis, polymerase chain reaction ("PCR") (e.g., Saiki et al., *Science*, 241:53, 1988; U.S. Pat. Nos. 4,683,195, 4,683,202, and 6,040,166; *PCR Protocols: A Guide to Methods and Applications*, Innis et al.,
25 eds., Academic Press, New York, 1990), reverse transcriptase polymerase chain reaction ("RT-PCR"), anchored PCR, rapid amplification of cDNA ends ("RACE") (e.g., Schaefer in *Gene Cloning and Analysis: Current Innovations*, Pages 99-115, 1997), ligase chain reaction ("LCR") (EP 320 308), one-sided PCR (Ohara et al., *Proc. Natl. Acad. Sci.*, 86:5673-5677, 1989), indexing methods (e.g., U.S. Pat. No. 5,508,169), *in situ* hybridization, differential
30 display (e.g., Liang et al., *Nucl. Acid. Res.*, 21:3269-3275, 1993; U.S. Pat. Nos. 5,262,311, 5,599,672 and 5,965,409; WO97/18454; Prashar and Weissman, *Proc. Natl. Acad. Sci.*,

93:659-663, and U.S. Pat. Nos. 6,010,850 and 5,712,126; Welsh et al., *Nucleic Acid Res.*, 20:4965-4970, 1992, and U.S. Pat. No. 5,487,985) and other RNA fingerprinting techniques, nucleic acid sequence based amplification ("NASBA") and other transcription based amplification systems (e.g., U.S. Pat. Nos. 5,409,818 and 5,554,527; WO 88/10315),
5 polynucleotide arrays (e.g., U.S. Pat. Nos. 5,143,854, 5,424,186; 5,700,637, 5,874,219, and 6,054,270; PCT WO 92/10092; PCT WO 90/15070), Qbeta Replicase (PCT/US87/00880), Strand Displacement Amplification ("SDA"), Repair Chain Reaction ("RCR"), nuclease protection assays, subtraction-based methods, Rapid-Scan™, etc. Additional useful methods include, but are not limited to, e.g., template-based amplification methods, competitive PCR
10 (e.g., U.S. Pat. No. 5,747,251), redox-based assays (e.g., U.S. Pat. No. 5,871,918), Taqman-based assays (e.g., Holland et al., *Proc. Natl. Acad. Sci.*, 88:7276-7280, 1991; U.S. Pat. Nos. 5,210,015 and 5,994,063), real-time fluorescence-based monitoring (e.g., U.S. Pat. 5,928,907), molecular energy transfer labels (e.g., U.S. Pat. Nos. 5,348,853, 5,532,129, 5,565,322, 6,030,787, and 6,117,635; Tyagi and Kramer, *Nature Biotech.*, 14:303-309,
15 1996). Any method suitable for single cell analysis of gene or protein expression can be used, including in situ hybridization, immunocytochemistry, MACS, FACS, flow cytometry, etc. For single cell assays, expression products can be measured using antibodies, PCR, or other types of nucleic acid amplification (e.g., Brady et al., *Methods Mol. & Cell. Biol.* 2, 17-25, 1990; Eberwine et al., 1992, *Proc. Natl. Acad. Sci.*, 89, 3010-3014, 1992; U.S. Pat. No.
20 5,723,290). These and other methods can be carried out conventionally, e.g., as described in the mentioned publications.

Many of such methods may require that the polynucleotide is labeled, or comprises a particular nucleotide type useful for detection. The present invention includes such modified polynucleotides that are necessary to carry out such methods. Thus, polynucleotides can be
25 DNA, RNA, DNA:RNA hybrids, PNA, etc., and can comprise any modification or substituent which is effective to achieve detection.

Detection can be desirable for a variety of different purposes, including research, diagnostic, prognostic, and forensic. For diagnostic purposes, it may be desirable to identify the presence or quantity of a polynucleotide sequence in a sample, where the sample is
30 obtained from tissue, cells, body fluids, etc. In a preferred method as described in more detail below, the present invention relates to a method of detecting a polynucleotide

comprising, contacting a target polynucleotide in a test sample with a polynucleotide probe under conditions effective to achieve hybridization between the target and probe; and detecting hybridization.

Any test sample in which it is desired to identify a polynucleotide or polypeptide thereof can be used, including, e.g., blood, urine, saliva, stool (for extracting nucleic acid, see, e.g., U.S. Pat. No. 6,177,251), swabs comprising tissue, biopsied tissue, tissue sections, cultured cells, etc.

Detection can be accomplished in combination with polynucleotide probes for other genes, e.g., genes which are expressed in other disease states, tissues, cells, such as brain, heart, kidney, spleen, thymus, liver, stomach, small intestine, colon, muscle, lung, testis, placenta, pituitary, thyroid, skin, adrenal gland, pancreas, salivary gland, uterus, ovary, prostate gland, peripheral blood cells (T-cells, lymphocytes, etc.), embryo, breast, fat, adult and embryonic stem cells, specific cell-types, such as endothelial, epithelial, myocytes, adipose, etc.

Polynucleotides can be used in wide range of methods and compositions, including for detecting, diagnosing, staging, grading, assessing, prognosticating, etc. diseases and disorders associated with polynucleotides, for monitoring or assessing therapeutic and/or preventative measures, in ordered arrays, etc. Any method of detecting genes and polynucleotides can be used; certainly, the present invention is not to be limited how such methods are implemented.

Along these lines, the present invention relates to methods of detecting polynucleotides in a sample comprising nucleic acid. Such methods can comprise one or more the following steps in any effective order, e.g., contacting said sample with a polynucleotide probe under conditions effective for said probe to hybridize specifically to nucleic acid in said sample, and detecting the presence or absence of probe hybridized to nucleic acid in said sample, wherein said probe is a polynucleotide which is SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58, a polynucleotide having, e.g., about 70%, 80%, 85%, 90%, 95%, 99%, or more sequence identity thereto, effective or specific fragments thereof, or complements thereto. The detection method can be applied to any sample, e.g., cultured primary, secondary, or established cell lines, tissue biopsy, blood, urine, stool, cerebral spinal fluid, and other bodily fluids, for any purpose.

-45-

Contacting the sample with probe can be carried out by any effective means in any effective environment. It can be accomplished in a solid, liquid, frozen, gaseous, amorphous, solidified, coagulated, colloid, etc., mixtures thereof, matrix. For instance, a probe in an aqueous medium can be contacted with a sample which is also in an aqueous medium, or
5 which is affixed to a solid matrix, or vice-versa.

Generally, as used throughout the specification, the term "effective conditions" means, e.g., the particular milieu in which the desired effect is achieved. Such a milieu, includes, e.g., appropriate buffers, oxidizing agents, reducing agents, pH, co-factors, temperature, ion concentrations, suitable age and/or stage of cell (such as, in particular part of
10 the cell cycle, or at a particular stage where particular genes are being expressed) where cells are being used, culture conditions (including substrate, oxygen, carbon dioxide, etc.). When hybridization is the chosen means of achieving detection, the probe and sample can be combined such that the resulting conditions are functional for said probe to hybridize specifically to nucleic acid in said sample.

15 The phrase "hybridize specifically" indicates that the hybridization between single-stranded polynucleotides is based on nucleotide sequence complementarity. The effective conditions are selected such that the probe hybridizes to a preselected and/or definite target nucleic acid in the sample. For instance, if detection of a polynucleotide set forth in SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58 is desired, a probe can be selected which
20 can hybridize to such target gene under high stringent conditions, without significant hybridization to other genes in the sample. To detect homologs of a polynucleotide set forth in SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58, the effective hybridization conditions can be less stringent, and/or the probe can comprise codon degeneracy, such that a homolog is detected in the sample.

25 As already mentioned, the methods can be carried out by any effective process, e.g., by Northern blot analysis, polymerase chain reaction (PCR), reverse transcriptase PCR, RACE PCR, *in situ* hybridization, etc., as indicated above. When PCR based techniques are used, two or more probes are generally used. One probe can be specific for a defined sequence which is characteristic of a selective polynucleotide, but the other probe can be
30 specific for the selective polynucleotide, or specific for a more general sequence, e.g., a sequence such as polyA which is characteristic of mRNA, a sequence which is specific for a

promoter, ribosome binding site, or other transcriptional features, a consensus sequence (e.g., representing a functional domain). For the former aspects, 5' and 3' probes (e.g., polyA, Kozak, etc.) are preferred which are capable of specifically hybridizing to the ends of transcripts. When PCR is utilized, the probes can also be referred to as "primers" in that they
5 can prime a DNA polymerase reaction.

In addition to testing for the presence or absence of polynucleotides, the present invention also relates to determining the amounts at which polynucleotides of the present invention are expressed in sample and determining the differential expression of such polynucleotides in samples. Such methods can involve substantially the same steps as
10 described above for presence/absence detection, e.g., contacting with probe, hybridizing, and detecting hybridized probe, but using more quantitative methods and/or comparisons to standards. The amount of hybridization between the probe and target can be determined by any suitable methods, e.g., PCR, RT-PCR, RACE PCR, Northern blot, polynucleotide microarrays, Rapid-Scan, etc., and includes both quantitative and qualitative measurements.

15 Methods of identifying polymorphisms, mutations, etc., of polynucleotides

Polynucleotides of the present invention can also be utilized to identify mutant alleles, SNPs, gene rearrangements and modifications, and other polymorphisms of the wild-type gene. Mutant alleles, polymorphisms, SNPs, etc., can be identified and isolated from
20 subjects with diseases that are known, or suspected to have, a genetic component. Identification of such genes can be carried out routinely (see, above for more guidance), e.g., using PCR, hybridization techniques, direct sequencing, mismatch reactions (see, e.g., above), RFLP analysis, SSCP (e.g., Orita et al., *Proc. Natl. Acad. Sci.*, 86:2766, 1992), etc., where a polynucleotide having a sequence selected from SEQ ID NOS 1, 11, 16, 25, 38, 43,
25 45, 47, 49, 51, and/or 58 is used as a probe. The selected mutant alleles, SNPs, polymorphisms, etc., can be used diagnostically to determine whether a subject has, or is susceptible to a disorder associated with polynucleotides, as well as to design therapies and predict the outcome of the disorder. Methods involve, e.g., diagnosing a disorder associated with polynucleotides or determining susceptibility to a disorder, comprising, detecting the
30 presence of a mutation in a gene represented by a polynucleotide selected from SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58. The detecting can be carried out by any

-47-

effective method, e.g., obtaining cells from a subject, determining the gene sequence or structure of a target gene (using, e.g., mRNA, cDNA, genomic DNA, etc), comparing the sequence or structure of the target gene to the structure of the normal gene, whereby a difference in sequence or structure indicates a mutation in the gene in the subject.

- 5 Polynucleotides can also be used to test for mutations, SNPs, polymorphisms, etc., e.g., using mismatch DNA repair technology as described in U.S. Pat. No. 5,683,877; U.S. Pat. No. 5,656,430; Wu et al., *Proc. Natl. Acad. Sci.*, 89:8779-8783, 1992.

The present invention also relates to methods of detecting polymorphisms in genes of the present invention, comprising, e.g., comparing the structure of: genomic DNA, mRNA,
10 cDNA, etc., with the structure of a polynucleotide set forth in SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, or 58. The methods can be carried out on a sample from any source, e.g., cells, tissues, body fluids, blood, urine, stool, hair, egg, sperm, cerebral spinal fluid, etc.

These methods can be implemented in many different ways. For example, "comparing the structure" steps include, but are not limited to, comparing restriction maps,
15 nucleotide sequences, amino acid sequences, RFLPs, Dnase sites, DNA methylation fingerprints (e.g., U.S. Pat. No. 6,214,556), protein cleavage sites, molecular weights, electrophoretic mobilities, charges, ion mobility, etc.. The term "structure" can refer to any physical characteristics or configurations which can be used to distinguish between nucleic acids and polypeptides. The methods and instruments used to accomplish the comparing step
20 depends upon the physical characteristics which are to be compared. Thus, various techniques are contemplated, including, e.g., sequencing machines (both amino acid and polynucleotide), electrophoresis, mass spectrometer (U.S. Pat. Nos. 6,093,541, 6,002,127), liquid chromatography, HPLC, etc.

To carry out such methods, "all or part" of the gene or polypeptide can be compared.
25 For example, if nucleotide sequencing is utilized, the entire gene can be sequenced, including promoter, introns, and exons, or only parts of it can be sequenced and compared, e.g., exon 1, exon 2, etc.

Mutagenesis

30 Mutated polynucleotide sequences of the present invention are useful for various purposes, e.g., to create mutations of the polypeptides they encode, to identify functional

regions of genomic DNA, to produce probes for screening libraries, etc. Mutagenesis can be carried out routinely according to any effective method, e.g., oligonucleotide-directed (Smith, M., *Ann. Rev. Genet.* 19:423-463, 1985), degenerate oligonucleotide-directed (Hill et al., *Method Enzymology*, 155:558-568, 1987), region-specific (Myers et al., *Science*, 229:242-246, 1985; Derbyshire et al., *Gene*, 46:145, 1986; Ner et al., *DNA*, 7:127, 1988), linker-scanning (McKnight and Kingsbury, *Science*, 217:316-324, 1982), directed using PCR, recursive ensemble mutagenesis (Arkin and Yourvan, *Proc. Natl. Acad. Sci.*, 89:7811-7815, 1992), random mutagenesis (e.g., U.S. Pat. Nos. 5,096,815; 5,198,346; and 5,223,409), site-directed mutagenesis (e.g., Walder et al., *Gene*, 42:133, 1986; Bauer et al., *Gene*, 37:73, 1985; Craik, *Bio Techniques*, January 1985, 12-19; Smith et al., *Genetic Engineering: Principles and Methods*, Plenum Press, 1981), phage display (e.g., Lowman et al., *Biochem.* 30:10832-10837, 1991; Ladner et al., U.S. Pat. No. 5,223,409; Huse, WIPO Publication WO 92/06204), etc. Desired sequences can also be produced by the assembly of target sequences using mutually priming oligonucleotides (Uhlmann, *Gene*, 71:29-40, 1988). For directed mutagenesis methods, analysis of the three-dimensional structure of the polynucleotides polypeptide can be used to guide and facilitate making mutants which effect polypeptide activity. Sites of substrate-enzyme interaction or other biological activities can also be determined by analysis of crystal structure as determined by such techniques as nuclear magnetic resonance, crystallography or photoaffinity labeling. See, for example, de Vos et al., *Science* 255:306-312, 1992; Smith et al., *J. Mol. Biol.* 224:899-904, 1992; Wlodaver et al., *FEBS Lett.* 309:59-64, 1992.

In addition, libraries of polynucleotides and fragments thereof can be used for screening and selection of polynucleotides variants. For instance, a library of coding sequences can be generated by treating a double-stranded DNA with a nuclease under conditions where the nicking occurs, e.g., only once per molecule, denaturing the double-stranded DNA, renaturing it to for double-stranded DNA that can include sense/antisense pairs from different nicked products, removing single-stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting DNAs into an expression vector. By this method, expression libraries can be made comprising "mutagenized" polynucleotides. The entire coding sequence or parts thereof can be used.

Polynucleotide expression, polypeptides produced thereby, and specific-binding partners thereto.

A polynucleotide according to the present invention can be expressed in a variety of different systems, in vitro and in vivo, according to the desired purpose. For example, a

5 polynucleotide can be inserted into an expression vector, introduced into a desired host, and cultured under conditions effective to achieve expression of a polypeptide coded for by the polynucleotide, to search for specific binding partners. Effective conditions include any culture conditions which are suitable for achieving production of the polypeptide by the host cell, including effective temperatures, pH, medium, additives to the media in which the host
10 cell is cultured (e.g., additives which amplify or induce expression such as butyrate, or methotrexate if the coding polynucleotide is adjacent to a dhfr gene), cycloheximide, cell densities, culture dishes, etc. A polynucleotide can be introduced into the cell by any effective method including, e.g., naked DNA, calcium phosphate precipitation, electroporation, injection, DEAE-Dextran mediated transfection, fusion with liposomes,
15 association with agents which enhance its uptake into cells, viral transfection. A cell into which a polynucleotide of the present invention has been introduced is a transformed host cell. The polynucleotide can be extrachromosomal or integrated into a chromosome(s) of the host cell. It can be stable or transient. An expression vector is selected for its compatibility with the host cell. Host cells include, mammalian cells, e.g., COS, CV1, BHK, CHO, HeLa,
20 LTK, NIH 3T3, cardiac or heart cells, such as W1 (Wang et al., *In vitro Cell. Dev.*, 27:63-74, 1991), MC29, cardiac fibroblasts (e.g., Wang et al., *Tiss Cell.*, 33:86-96, 2001), cardiac microvascular endothelial cells (e.g. Jollow et al., *Transplantation*, 68:430-439, 1999), T/G HA-VSMC (CRL-1999), H9c2(2-1) (CRL-1446), P19 (CRL-1825), CNS neural stem cells (e.g., U.S. Pat. No. 6,103,530), IMR-32, A172 (ATCC CRL-1620), T98G (ATCC CRL-
25 1690), CCF-STTG1 (ATCC CRL-1718), DBTRG-05MG (ATCC CRL-2020), PFSK-1 (ATCC CRL-2060), SK-N-AS and other SK cell lines (ATCC CRL-2137), CHP-212 (ATCC CRL-2273), RG2 (ATCC CRL-2433), HCN-2 (ATCC CRL-10742), U-87 MG and other U MG cell lines (ATCC HTB-14), D283 Med (ATCC HTB-185), PC12, Neuro-2a (ATCC CCL-131), muscle cells lines, such as RD (CCL-136), G-7, G-8, C2C12, established and
30 primary brain, heart, or muscle cells, G-402 (ATCC CRL-1440), ACHN (ATCC CRL-1611), Vero (ATCC CCL-81), 786-O (ATCC CRL-1932), 769-P (ATCC CRL-1933), CCD 1103

KIDTr (ATCC CRL-2304), CCD 1105 KIDTr (ATCC CRL-2305), Hs 835.T (ATCC CRL-7569), Hs 926.T (ATCC CRL-7678), Caki-1 (ATCC HTB-46), Caki-2 (ATCC HTB-47), SW 839 (ATCC HTB-49), LLC-MK2 (ATCC CCL-7), BHK-21 (ATCC CCL-10), MDBK, CV-1, (ATCC CRL-1573), KNRK (ATCC CRL-1569), NRK-49F (ATCC CRL-1570), A-704
5 (ATCC HTB-45), and other established and primary kidney lines, CNS neural stem cells (e.g., U.S. Pat. No. 6,103,530), IMR-32, A172 (ATCC CRL-1620), T98G (ATCC CRL-1690), CCF-STTG1 (ATCC CRL-1718), DBTRG-05MG (ATCC CRL-2020), PFSK-1 (ATCC CRL-2060), SK-N-AS and other SK cell lines (ATCC CRL-2137), CHP-212 (ATCC CRL-2273), RG2 (ATCC CRL-2433), HCN-2 (ATCC CRL-10742), U-87 MG and other U
10 MG cell lines (ATCC HTB-14), D283 Med (ATCC HTB-185), PC12, Neuro-2a (ATCC CCL-131), and other established and primary brain cell lines, Hep G2 (ATCC NO. HB-8065), SK-HEP-1 (ATCC NO HTB-52), H2.35 (ATCC NO CRL-1995), CD-1 (ATC NO CRL-2254), C3A (ATCC NO CRL-10741), FL83B (ATCC NO CRL-2390), WRL 68 (ATCC NO CL-48), Hep 3B (ATCC NO HB-8064), insect cells, such as Sf9 (*S. frugipeda*)
15 and *Drosophila*, bacteria, such as *E. coli*, *Streptococcus*, *bacillus*, yeast, such as *Sacharomyces*, *S. cerevisiae*, fungal cells, plant cells, embryonic or adult stem cells (e.g., mammalian, such as mouse or human).

Expression control sequences are similarly selected for host compatibility and a desired purpose, e.g., high copy number, high amounts, induction, amplification, controlled
20 expression. Other sequences which can be employed include enhancers such as from SV40, CMV, RSV, inducible promoters, cell-type specific elements, or sequences which allow selective or specific cell expression. Promoters that can be used to drive its expression, include, e.g., the endogenous promoter, MMTV, SV40, trp, lac, tac, or T7 promoters for bacterial hosts; or alpha factor, alcohol oxidase, or PGH promoters for yeast. RNA
25 promoters can be used to produced RNA transcripts, such as T7 or SP6. See, e.g., Melton et al., *Polynucleotide Res.*, 12(18):7035-7056, 1984; Dunn and Studier. *J. Mol. Bio.*, 166:477-435, 1984; U.S. Pat. No. 5,891,636; Studier et al., *Gene Expression Technology, Methods in Enzymology*, 85:60-89, 1987. In addition, as discussed above, translational signals (including in-frame insertions) can be included.

30 When a polynucleotide is expressed as a heterologous gene in a transfected cell line, the gene is introduced into a cell as described above, under effective conditions in which the

gene is expressed. The term "heterologous" means that the gene has been introduced into the cell line by the "hand-of-man." Introduction of a gene into a cell line is discussed above. The transfected (or transformed) cell expressing the gene can be lysed or the cell line can be used intact.

5 For expression and other purposes, a polynucleotide can contain codons found in a naturally-occurring gene, transcript, or cDNA, for example, e.g., as set forth in SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58, or it can contain degenerate codons coding for the same amino acid sequences. For instance, it may be desirable to change the codons in the sequence to optimize the sequence for expression in a desired host. See, e.g., U.S. Pat. Nos.
10 5,567,600 and 5,567,862.

A polypeptide according to the present invention can be recovered from natural sources, transformed host cells (culture medium or cells) according to the usual methods, Another approach is express the polypeptide recombinantly with an affinity tag (Flag epitope, HA epitope, myc epitope, 6xHis, maltose binding protein, chitinase, etc) and then purify by
15 anti-tag antibody-conjugated affinity chromatography.

The present invention also relates to polypeptides, e.g., an isolated human polypeptide comprising or having the amino acid sequence set forth in SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58, an isolated human polypeptide comprising an amino acid sequence having 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more amino acid
20 sequence identity to the amino acid sequence set forth in SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58, and optionally having one or more activities. Fragments can also be used, e.g., to produce antibodies or other immune responses, as competitors to any activity. These fragments can be referred to as being "specific for" polynucleotides. The latter phrase, as already defined, indicates that the peptides are characteristic of polynucleotides, and that
25 the defined sequences are substantially absent from all other protein types. Such polypeptides can be of any size which is necessary to confer specificity, e.g., 5, 8, 10, 12, 15, 20, etc.

The present invention also relates to specific-binding partners. These include antibodies which are specific for polypeptides encoded by polynucleotides of the present
30 invention, as well as other binding-partners which interact with polynucleotides and polypeptides of the present invention. Protein-protein interactions between polynucleotides

and other polypeptides and binding partners can be identified using any suitable methods, e.g., protein binding assays (e.g., filtration assays, chromatography, etc.), yeast two-hybrid system (Fields and Song, *Nature*, 340: 245-247, 1989), protein arrays, gel-shift assays, FRET (fluorescence resonance energy transfer) assays, etc. Nucleic acid interactions (e.g., protein-DNA or protein-RNA) can be assessed using gel-shift assays, e.g., as carried out in U.S. Pat. No. 6,333,407 and 5,789,538.

Antibodies, e.g., polyclonal, monoclonal, recombinant, chimeric, humanized, single-chain, Fab, and fragments thereof, can be prepared according to any desired method. See, also, screening recombinant immunoglobulin libraries (e.g., Orlandi et al., *Proc. Natl. Acad. Sci.*, 86:3833-3837, 1989; Huse et al., *Science*, 256:1275-1281, 1989); in vitro stimulation of lymphocyte populations; Winter and Milstein, *Nature*, 349: 293-299, 1991. The antibodies can be IgM, IgG, subtypes, IgG2a, IgG1, etc. Antibodies, and immune responses, can also be generated by administering naked DNA See, e.g., U.S. Pat. Nos. 5,703,055; 5,589,466; 5,580,859. Antibodies can be used from any source, including, goat, rabbit, mouse, chicken (e.g., IgY; see, Duan, W0/029444 for methods of making antibodies in avian hosts, and harvesting the antibodies from the eggs). An antibody specific for a polypeptide means that the antibody recognizes a defined sequence of amino acids within or including the polypeptide. Other specific binding partners include, e.g., aptamers and PNA. Antibodies can be prepared against specific epitopes or domains of polynucleotides.

Any form or type of antibody can be prepared and used. For example, antibodies can be humanized, e.g., where they are to be used therapeutically. Another form of an antibody fragment is a peptide coding for a single complementarity-determining region (CDR). CDR peptides ("minimal recognition units") can be obtained by constructing genes encoding the CDR of an antibody of interest. The term "antibody" as used herein includes intact molecules as well as fragments thereof, such as Fab, F(ab')₂, and Fv which are capable of binding to an epitopic determinant. Such antibody fragments retain some ability to selectively bind with its antigen or receptor. The term "epitope" refers to an antigenic determinant on an antigen to which the paratope of an antibody binds. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three dimensional structural characteristics, as well as specific

charge characteristics. Antibodies can be prepared against specific epitopes or polypeptide domains.

Antibodies which bind to polynucleotides polypeptides of the present invention can be prepared using an intact polypeptide or fragments containing small peptides of interest as the immunizing antigen. For example, it may be desirable to produce antibodies that specifically bind to the N- or C-terminal domains of polynucleotides. The polypeptide or peptide used to immunize an animal which is derived from translated cDNA or chemically synthesized which can be conjugated to a carrier protein, if desired. Anti-idiotypic technology can also be used to produce invention monoclonal antibodies which mimic an epitope.

Methods of detecting polypeptides

Polypeptides coded for by polynucleotides of the present invention can be detected, visualized, determined, quantitated, etc. according to any effective method. useful methods include, e.g., but are not limited to, immunoassays, RIA (radioimmunoassay), ELISA, (enzyme-linked-immunosorbent assay), immunofluorescence, flow cytometry, histology, electron microscopy, light microscopy, in situ assays, immunoprecipitation, Western blot.

Immunoassays may be carried in liquid or on biological support. For instance, a sample (e.g., blood, stool, urine, cells, tissue, cerebral spinal fluid, body fluids, etc.) can be brought in contact with and immobilized onto a solid phase support or carrier such as nitrocellulose, or other solid support that is capable of immobilizing cells, cell particles or soluble proteins. The support may then be washed with suitable buffers followed by treatment with the detectably labeled polynucleotides specific antibody. The solid phase support can then be washed with a buffer a second time to remove unbound antibody. The amount of bound label on solid support may then be detected by conventional means.

A "solid phase support or carrier" includes any support capable of binding an antigen, antibody, or other specific binding partner. Supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, and magnetite. A support material can have any structural or physical configuration. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the

surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads

One of the many ways in which gene peptide-specific antibody can be detectably labeled is by linking it to an enzyme and using it in an enzyme immunoassay (EIA). See, e.g., Voller, A., "The Enzyme Linked Immunosorbent Assay (ELISA)," 1978, Diagnostic Horizons 2, 1-7, Microbiological Associates Quarterly Publication, Walkersville, Md.); Voller, A. et al., 1978, J. Clin. Pathol. 31, 507-520; Butler, J. E., 1981, Meth. Enzymol. 73, 482-523; Maggio, E. (ed.), 1980, Enzyme Immunoassay, CRC Press, Boca Raton, Fla.. The enzyme which is bound to the antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety that can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes that can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, .alpha.-glycerophosphate dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, .beta.-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods that employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Detection may also be accomplished using any of a variety of other immunoassays. For example, by radioactively labeling the antibodies or antibody fragments, it is possible to detect polynucleotides peptides through the use of a radioimmunoassay (RIA). See, e.g., Weintraub, B., Principles of Radioimmunoassays, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986. The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin,

-55-

allophycocyanin, o-phthaldehyde and fluorescamine. The antibody can also be detectably labeled using fluorescence emitting metals such as those in the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentaacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

5 The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of useful chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester.

10 Likewise, a bioluminescent compound may be used to label the antibody of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Important bioluminescent compounds for purposes of labeling are luciferin,
15 luciferase and aequorin.

Diagnostic

 The present invention also relates to methods and compositions for diagnosing a disorder, or determining susceptibility to a disorder, using polynucleotides, polypeptides, and
20 specific-binding partners of the present invention to detect, assess, determine, etc., polynucleotides of the present invention. In such methods, the gene can serve as a marker for the disorder, e.g., where the gene, when mutant, is a direct cause of the disorder; where the gene is affected by another gene(s) which is directly responsible for the disorder, e.g., when the gene is part of the same signaling pathway as the directly responsible gene; and,
25 where the gene is chromosomally linked to the gene(s) directly responsible for the disorder, and segregates with it. Many other situations are possible. To detect, assess, determine, etc., a probe specific for the gene can be employed as described above and below. Any method of detecting and/or assessing the gene can be used, including detecting expression of the gene using polynucleotides, antibodies, or other specific-binding partners.

30 The present invention relates to methods of diagnosing a disorder associated with a polynucleotide of the present invention, or determining a subject's susceptibility to such

disorder, comprising, e.g., assessing the expression of polynucleotide of the present invention in a tissue sample comprising tissue or cells suspected of having the disorder. The phrase “diagnosing” indicates that it is determined whether the sample has the disorder. A “disorder” means, e.g., any abnormal condition as in a disease or malady. “Determining a subject’s susceptibility to a disease or disorder” indicates that the subject is assessed for whether s/he is predisposed to get such a disease or disorder, where the predisposition is indicated by abnormal expression of the gene (e.g., gene mutation, gene expression pattern is not normal, etc.). Predisposition or susceptibility to a disease may result when a such disease is influenced by epigenetic, environmental, etc., factors. Diagnosing includes prenatal screening where samples from the fetus or embryo (e.g., via amniocentesis or CV sampling) are analyzed for the expression of the gene.

By the phrase “assessing expression of a gene or polynucleotide of the present invention,” it is meant that the functional status of the gene is evaluated. This includes, but is not limited to, measuring expression levels of said gene, determining the genomic structure of said gene, determining the mRNA structure of transcripts from said gene, or measuring the expression levels of polypeptide coded for by said gene. Thus, the term “assessing expression” includes evaluating the all aspects of the transcriptional and translational machinery of the gene. For instance, if a promoter defect causes, or is suspected of causing, the disorder, then a sample can be evaluated (i.e., “assessed”) by looking (e.g., sequencing or restriction mapping) at the promoter sequence in the gene, by detecting transcription products (e.g., RNA), by detecting translation product (e.g., polypeptide). Any measure of whether the gene is functional can be used, including, polypeptide, polynucleotide, and functional assays for the gene’s biological activity.

In making the assessment, it can be useful to compare the results to a normal gene, e.g., a gene which is not associated with the disorder. The nature of the comparison can be determined routinely, depending upon how the assessing is accomplished. If, for example, the mRNA levels of a sample is detected, then the mRNA levels of a normal can serve as a comparison, or a gene which is known not to be affected by the disorder. Methods of detecting mRNA are well known, and discussed above, e.g., but not limited to, Northern blot analysis, polymerase chain reaction (PCR), reverse transcriptase PCR, RACE PCR, etc. Similarly, if polypeptide production is used to evaluate the gene, then the polypeptide in a

-57-

normal tissue sample can be used as a comparison, or, polypeptide from a different gene whose expression is known not to be affected by the disorder. These are only examples of how such a method could be carried out.

The present invention relates to methods of identifying a genetic basis for a disease or disease-susceptibility, comprising, e.g., determining the association of a disease or disease-susceptibility with a gene of the present invention. An association between a disease or disease-susceptibility and nucleotide sequence includes, e.g., establishing (or finding) a correlation (or relationship) between a DNA marker (e.g., gene, VNTR, polymorphism, EST, etc.) and a particular disease state. Once a relationship is identified, the DNA marker can be utilized in diagnostic tests and as a drug target. Any region of the gene can be used as a source of the DNA marker, exons, introns, intergenic regions, etc.

Human linkage maps can be constructed to establish a relationship between a gene and a disease or condition. Typically, polymorphic molecular markers (e.g., STRP's, SNP's, RFLP's, VNTR's) are identified within the region, linkage and map distance between the markers is then established, and then linkage is established between phenotype and the various individual molecular markers. Maps can be produced for an individual family, selected populations, patient populations, etc. In general, these methods involve identifying a marker associated with the disease (e.g., identifying a polymorphism in a family which is linked to the disease) and then analyzing the surrounding DNA to identify the gene responsible for the phenotype. See, e.g., Kruglyak et al., *Am. J. Hum. Genet.*, 58, 1347-1363, 1996; Matisse et al., *Nat. Genet.*, 6(4):384-90, 1994.

Assessing the effects of therapeutic and preventative interventions (e.g., administration of a drug, chemotherapy, radiation, etc.) is a major effort in drug discovery, clinical medicine, and pharmacogenomics. The evaluation of therapeutic and preventative measures, whether experimental or already in clinical use, has broad applicability, e.g., in clinical trials, for monitoring the status of a patient, for analyzing and assessing animal models, and in any scenario involving disease treatment and prevention. Analyzing the expression profiles of polynucleotides of the present invention can be utilized as a parameter by which interventions are judged and measured. Treatment of a disorder can change the expression profile in some manner which is prognostic or indicative of the drug's effect on it. Changes in the profile can indicate, e.g., drug toxicity, return to a normal level, etc.

Accordingly, the present invention also relates to methods of monitoring or assessing a therapeutic or preventative measure (e.g., chemotherapy, radiation, anti-neoplastic drugs, antibodies, etc.) in a subject having a disorder, or, susceptible to such a disorder, comprising, e.g., detecting the expression levels of polynucleotides. A subject can be a cell-based assay system, non-human animal model, human patient, etc. Detecting can be accomplished as described for the methods above and below. By "therapeutic or preventative intervention," it is meant, e.g., a drug administered to a patient, surgery, radiation, chemotherapy, and other measures taken to prevent, treat, or diagnose a disorder. Expression can be assessed in any sample comprising any tissue or cell type, body fluid, etc., as discussed for other methods of the present invention.

The present invention also relates to methods of using polynucleotides binding partners, such as antibodies, to deliver active agents to a tissue for a variety of different purposes, including, e.g., for diagnostic, therapeutic, and research purposes. Methods can involve delivering or administering an active agent to the target tissue, comprising, e.g., administering to a subject in need thereof, an effective amount of an active agent coupled to a binding partner specific for a polypeptide of the present invention, wherein said binding partner is effective to deliver said active agent specifically to the target tissue.

Any type of active agent can be used in combination with polynucleotides, including, therapeutic, cytotoxic, cytostatic, chemotherapeutic, anti-neoplastic, anti-proliferative, anti-biotic, etc., agents. A chemotherapeutic agent can be, e.g., DNA-interactive agent, alkylating agent, antimetabolite, tubulin-interactive agent, hormonal agent, hydroxyurea, Cisplatin, Cyclophosphamide, Altretamine, Bleomycin, Dactinomycin, Doxorubicin, Etoposide, Teniposide, paclitaxel, cytoxan, 2-methoxycarbonylaminobenzimidazole, Plicamycin, Methotrexate, Fluorouracil, Fluorodeoxyuridin, CB3717, Azacitidine, Floxuridine, Mercaptopurine, 6-Thioguanine, Pentostatin, Cytarabine, Fludarabine, etc. Agents can also be contrast agents useful in imaging technology, e.g., X-ray, CT, CAT, MRI, ultrasound, PET, SPECT, and scintographic.

An active agent can be associated in any manner with a polynucleotides binding partner which is effective to achieve its delivery specifically to the target. Specific delivery or targeting indicates that the agent is provided to the target, without being substantially provided to other tissues. This is useful especially where an agent is toxic, and specific

targeting enables the majority of the toxicity to be aimed at the target tissue, with as small as possible effect on other tissues in the body. The association of the active agent and the binding partner ("coupling") can be direct, e.g., through chemical bonds between the binding partner and the agent, or, via a linking agent, or the association can be less direct, e.g., where
5 the active agent is in a liposome, or other carrier, and the binding partner is associated with the liposome surface. In such case, the binding partner can be oriented in such a way that it is able to bind to polypeptide on the cell surface. Methods for delivery of DNA via a cell-surface receptor is described, e.g., in U.S. Pat. No. 6,339,139.

10 Identifying agent methods

The present invention also relates to methods of identifying agents, and the agents themselves, which modulate polynucleotides and polypeptides of the present invention. These agents can be used to modulate the biological activity of the polypeptide encoded for by the gene, or the gene, itself. Agents which regulate the gene or its product are useful in
15 variety of different environments, including as medicinal agents to treat or prevent disorders associated with polynucleotides of the present invention and as research reagents to modify the function of tissues and cell.

Methods of identifying agents generally comprise steps in which an agent is placed in contact with the gene, its transcription product, its translation product, or other target, and
20 then a determination is performed to assess whether the agent "modulates" the target. The specific method utilized will depend upon a number of factors, including, e.g., the target (i.e., is it the gene or polypeptide encoded by it), the environment (e.g., in vitro or in vivo), the composition of the agent, etc.

For modulating the expression of a gene or polynucleotide, a method can comprise, in
25 any effective order, one or more of the following steps, e.g., contacting a polynucleotide or gene (e.g., in a cell population) with a test agent under conditions effective for said test agent to modulate its expression, and determining whether said test agent modulates it. An agent can modulate expression of polynucleotides at any level, including transcription (e.g., by modulating the promoter), translation, and/or perdurance of the nucleic acid (e.g.,
30 degradation, stability, etc.) in the cell.

For modulating the biological activity of a polypeptide, a method can comprise, in any effective order, one or more of the following steps, e.g., contacting a polypeptide (e.g., in a cell, lysate, or isolated) with a test agent under conditions effective for said test agent to modulate the biological activity of said polypeptide, and determining whether said test agent
5 modulates said biological activity.

Contacting polynucleotides with the test agent can be accomplished by any suitable method and/or means that places the agent in a position to functionally control expression or biological activity. Functional control indicates that the agent can exert its physiological effect on polynucleotides through whatever mechanism it works. The choice of the method
10 and/or means can depend upon the nature of the agent and the condition and type of environment in which the polynucleotides is presented, e.g., lysate, isolated, or in a cell population (such as, *in vivo*, *in vitro*, organ explants, etc.). For instance, if the cell population is an *in vitro* cell culture, the agent can be contacted with the cells by adding it directly into the culture medium. If the agent cannot dissolve readily in an aqueous medium, it can be
15 incorporated into liposomes, or another lipophilic carrier, and then administered to the cell culture. Contact can also be facilitated by incorporation of agent with carriers and delivery molecules and complexes, by injection, by infusion, etc.

Agents can be directed to, or targeted to, any part of the polypeptide which is effective for modulating it. For example, agents, such as antibodies and small molecules, can
20 be targeted to cell-surface, exposed, extracellular, ligand binding, functional, etc., domains of the polypeptide. Agents can also be directed to intracellular regions and domains, e.g., regions where the polypeptide couples or interacts with intracellular or intramembrane binding partners.

After the agent has been administered in such a way that it can gain access, it can be
25 determined whether the test agent modulates expression or biological activity. Modulation can be of any type, quality, or quantity, e.g., increase, facilitate, enhance, up-regulate, stimulate, activate, amplify, augment, induce, decrease, down-regulate, diminish, lessen, reduce, etc. The modulatory quantity can also encompass any value, e.g., 1%, 5%, 10%, 50%, 75%, 1-fold, 2-fold, 5-fold, 10-fold, 100-fold, etc. To modulate expression means, e.g.,
30 that the test agent has an effect on its expression, e.g., to effect the amount of transcription, to effect RNA splicing, to effect translation of the RNA into polypeptide, to effect RNA or

-61-

polypeptide stability, to effect polyadenylation or other processing of the RNA, to effect post-transcriptional or post-translational processing, etc. To modulate biological activity means, e.g., that a functional activity of the polypeptide is changed in comparison to its normal activity in the absence of the agent. This effect includes, increase, decrease, block, inhibit, enhance, etc.

A test agent can be of any molecular composition, e.g., chemical compounds, biomolecules, such as polypeptides, lipids, nucleic acids (e.g., antisense to a polynucleotide sequence selected from SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58), carbohydrates, antibodies, ribozymes, double-stranded RNA, aptamers, etc. For example, if a polypeptide to be modulated is a cell-surface molecule, a test agent can be an antibody that specifically recognizes it and, e.g., causes the polypeptide to be internalized, leading to its down regulation on the surface of the cell. Such an effect does not have to be permanent, but can require the presence of the antibody to continue the down-regulatory effect. Antibodies can also be used to modulate the biological activity of a polypeptide in a lysate or other cell-free form.

Additional cell-based test systems suitable for the analysis of GPCR polypeptides are summarized in Marchese et al. (1999, Trends in Pharmacol. Sci. 20: 370-375) and comprise so-called "ligand screening assays." For example in yeast cells the pheromon receptor can be replaced by a GPCR according to the invention. The effect of test substances on the receptor can be determined upon modulation of histidine synthesis, i.e. by growing in histidine-free medium. In addition using cells transfected with nucleic acids according to the invention it can be analyzed whether test substances mediate translocation of a detectable arrestins, for example of a arrestin-GFP-fusion protein. Moreover, it can be analyzed whether test substances mediate GPCR-mediated dispersion or aggregation of *Xenopus laevis* melanophores. Another test system utilizes the universal adapter G-protein G α_{16} , which mobilizes Ca^{2+} . Other screening test systems are described in Lemer et al., supra; WO96/41169; U.S. Pat. No. 5,482,835; WO99/06535; EP 0 939 902; WO99/66326; WO98/34948; EP 0 863 214; U.S. Pat. No. 5,882,944 and U.S. Pat. No. 5,891,641.

Therapeutics

Selective polynucleotides, polypeptides, and specific-binding partners thereto, can be

utilized in therapeutic applications. Useful methods include, but are not limited to, immunotherapy (e.g., using specific-binding partners to polypeptides), vaccination (e.g., using a selective polypeptide or a naked DNA encoding such polypeptide), protein or polypeptide replacement therapy, gene therapy (e.g., germ-line correction, antisense), etc.

5 Various immunotherapeutic approaches can be used. For instance, unlabeled antibody that specifically recognizes a tissue-specific antigen can be used to stimulate the body to destroy or attack a cancer or other diseased tissue, to cause down-regulation, to produce complement-mediated lysis, to inhibit cell growth, etc., of target cells which display the antigen, e.g., analogously to how c-erbB-2 antibodies are used to treat breast cancer. In
10 addition, antibody can be labeled or conjugated to enhance its deleterious effect, e.g., with radionuclides and other energy emitting entities, toxins, such as ricin, exotoxin A (ETA), and diphtheria, cytotoxic or cytostatic agents, immunomodulators, chemotherapeutic agents, etc. See, e.g., U.S. Pat. No. 6,107,090.

 An antibody or other specific-binding partner can be conjugated to a second molecule,
15 such as a cytotoxic agent, and used for targeting the second molecule to a tissue-antigen positive cell (Vitetta, E. S. et al., 1993, Immunotoxin therapy, in DeVita, Jr., V. T. et al., eds, Cancer: Principles and Practice of Oncology, 4th ed., J. B. Lippincott Co., Philadelphia, 2624-2636). Examples of cytotoxic agents include, but are not limited to, antimetabolites, alkylating agents, anthracyclines, antibiotics, anti-mitotic agents, radioisotopes and
20 chemotherapeutic agents. Further examples of cytotoxic agents include, but are not limited to ricin, doxorubicin, daunorubicin, taxol, ethidium bromide, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicine, dihydroxy anthracin dione, actinomycin D, 1-dehydrotestosterone, diphtheria toxin, Pseudomonas exotoxin (PE) A, PE40, abrin, elongation factor-2 and glucocorticoid. Techniques for conjugating therapeutic agents to antibodies are
25 well.

 In addition to immunotherapy, polynucleotides and polypeptides can be used as targets for non-immunotherapeutic applications, e.g., using compounds which interfere with function, expression (e.g., antisense as a therapeutic agent), assembly, etc. RNA interference can be used in vitro and in vivo to silence polynucleotides when its expression contributes to
30 a disease (but also for other purposes, e.g., to identify the gene's function to change a developmental pathway of a cell, etc.). See, e.g., Sharp and Zamore, *Science*, 287:2431-

2433, 2001; Grishok et al., *Science*, 287:2494, 2001.

Delivery of therapeutic agents can be achieved according to any effective method, including, liposomes, viruses, plasmid vectors, bacterial delivery systems, orally, systemically, etc. Therapeutic agents of the present invention can be administered in any form by any effective route, including, e.g., oral, parenteral, enteral, intraperitoneal, topical, transdermal (e.g., using any standard patch), intravenously, ophthalmic, nasally, local, non-oral, such as aerosol, inhalation, subcutaneous, intramuscular, buccal, sublingual, rectal, vaginal, intra-arterial, and intrathecal, etc. They can be administered alone, or in combination with any ingredient(s), active or inactive.

In addition to therapeutics, *per se*, the present invention also relates to methods of treating a disease showing altered expression of a polynucleotide or polypeptide of the present invention, comprising, e.g., administering to a subject in need thereof a therapeutic agent which is effective for regulating expression of said polynucleotide or polypeptide which is effective in treating said disease. The term "treating" is used conventionally, e.g., the management or care of a subject for the purpose of combating, alleviating, reducing, relieving, improving the condition of, etc., of a disease or disorder. By the phrase "altered expression," it is meant that the disease is associated with a mutation in the gene, or any modification to the gene (or corresponding product) which affects its normal function. Thus, expression of polynucleotides refers to, e.g., transcription, translation, splicing, stability of the mRNA or protein product, activity of the gene product, differential expression, etc.

Any agent which "treats" the disease can be used. Such an agent can be one which regulates the expression of the polynucleotides. Expression refers to the same acts already mentioned, e.g. transcription, translation, splicing, stability of the mRNA or protein product, activity of the gene product, differential expression, etc. For instance, if the condition was a result of a complete deficiency of the gene product, administration of gene product to a patient would be said to treat the disease and regulate the gene's expression. Many other possible situations are possible, e.g., where the gene is aberrantly expressed, and the therapeutic agent regulates the aberrant expression by restoring its normal expression pattern.

Antisense

Antisense polynucleotide (e.g., RNA) can also be prepared from a polynucleotide according to the present invention, preferably an anti-sense to a sequence of SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58. Antisense polynucleotide can be used in various ways, such as to regulate or modulate expression of the polypeptides they encode, e.g., inhibit
5 their expression, for in situ hybridization, for therapeutic purposes, for making targeted mutations (in vivo, triplex, etc.) etc. For guidance on administering and designing anti-sense, see, e.g., U.S. Pat. Nos. 6,200,960, 6,200,807, 6,197,584, 6,190,869, 6,190,661, 6,187,587, 6,168,950, 6,153,595, 6,150,162, 6,133,246, 6,117,847, 6,096,722, 6,087,343, 6,040,296, 6,005,095, 5,998,383, 5,994,230, 5,891,725, 5,885,970, and 5,840,708. An antisense
10 polynucleotides can be operably linked to an expression control sequence. A total length of about 35 bp can be used in cell culture with cationic liposomes to facilitate cellular uptake, but for *in vivo* use, preferably shorter oligonucleotides are administered, e.g. 25 nucleotides.

Antisense polynucleotides can comprise modified, nonnaturally-occurring nucleotides and linkages between the nucleotides (e.g., modification of the phosphate-sugar backbone; methyl phosphonate, phosphorothioate, or phosphorodithioate linkages; and 2'-O-methyl
15 ribose sugar units), e.g., to enhance in vivo or in vitro stability, to confer nuclease resistance, to modulate uptake, to modulate cellular distribution and compartmentalization, etc. Any effective nucleotide or modification can be used, including those already mentioned, as known in the art, etc., e.g., disclosed in U.S. Pat. Nos. 6,133,438; 6,127,533; 6,124,445; 6,121,437; 5,218,103 (e.g., nucleoside thiophosphoramidites); 4,973,679; Sproat et al., "2'-O-Methyloligoribonucleotides: synthesis and applications," *Oligonucleotides and Analogs A Practical Approach*, Eckstein (ed.), IRL Press, Oxford, 1991, 49-86; Iribarren et al., "2'-O-Alkyl Oligoribonucleotides as Antisense Probes," *Proc. Natl. Acad. Sci. USA*, 1990, 87, 7747-7751; Cotton et al., "2'-O-methyl, 2'-O-ethyl oligoribonucleotides and phosphorothioate
20 oligodeoxyribonucleotides as inhibitors of the in vitro U7 snRNP-dependent mRNA processing event," *Nucl. Acids Res.*, 1991, 19, 2629-2635.

Arrays

The present invention also relates to an ordered array of polynucleotide probes and
30 specific-binding partners (e.g., antibodies) for detecting the expression of polynucleotides or polypeptides in a sample, comprising, e.g., one or more polynucleotide probes or specific

-65-

binding partners associated with a solid support or in separate receptacles, wherein each probe is specific for polynucleotides, and the probes comprise a nucleotide sequence of SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58 which is specific for said gene, a nucleotide sequence having sequence identity to SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58 which is specific for said gene or polynucleotide, or complements thereto, or a specific-binding partner which is specific for polynucleotides.

The phrase "ordered array" indicates that the probes (polynucleotides, binding-partners, polypeptides, etc.) are arranged in an identifiable or position-addressable pattern, e.g., such as the arrays disclosed in U.S. Pat. Nos. 6,156,501, 6,077,673, 6,054,270, 5,723,320, 5,700,637, WO09919711, WO00023803. The probes are associated with the solid support in any effective way. For instance, the probes can be bound to the solid support, either by polymerizing the probes on the substrate, or by attaching a probe to the substrate. Association can be, covalent, electrostatic, noncovalent, hydrophobic, hydrophilic, noncovalent, coordination, adsorbed, absorbed, polar, etc. When fibers or hollow filaments are utilized for the array, the probes can fill the hollow orifice, be absorbed into the solid filament, be attached to the surface of the orifice, etc. Probes can be of any effective size, sequence identity, composition, etc., as already discussed.

Transgenic animals

The present invention also relates to transgenic animals comprising polynucleotides genes, and homologs thereof. (Methods of making transgenic animals, and associated recombinant technology, can be accomplished conventionally, e.g., as described in *Transgenic Animal Technology*, Pinkert et al., 2nd Edition, Academic Press, 2002.) Such genes, as discussed in more detail below, include, but are not limited to, functionally-disrupted genes, mutated genes, ectopically or selectively-expressed genes, inducible or regulatable genes, etc. These transgenic animals can be produced according to any suitable technique or method, including homologous recombination, mutagenesis (e.g., ENU, Rathkolb et al., *Exp. Physiol.*, 85(6):635-644, 2000), and the tetracycline-regulated gene expression system (e.g., U.S. Pat. No. 6,242,667). The term "gene" as used herein includes any part of a gene, i.e., regulatory sequences, promoters, enhancers, exons, introns, coding sequences, etc. The polynucleotides nucleic acid present in the construct or transgene can be

naturally-occurring wild-type, polymorphic, or mutated. Where the animal is a non-human animal, its homolog can be used instead. Transgenic animals can be susceptible to any of the diseases and disorders mentioned herein, e.g., as described more particularly under the descriptions of each gene.

5 Along these lines, polynucleotides of the present invention can be used to create transgenic animals, e.g. a non-human animal, comprising at least one cell whose genome comprises a functional disruption of a polynucleotide of the present invention, or a homolog thereof (e.g., a mouse homolog when a mouse is used). By the phrases “functional disruption” or “functionally disrupted,” it is meant that the gene does not express a
10 biologically-active product. It can be substantially deficient in at least one functional activity coded for by the gene. Expression of a polypeptide can be substantially absent, i.e., essentially undetectable amounts are made. However, polypeptide can also be made, but which is deficient in activity, e.g., where only an amino-terminal portion of the gene product is produced.

15 The transgenic animal can comprise one or more cells. When substantially all its cells contain the engineered gene, it can be referred to as a transgenic animal “whose genome comprises” the engineered gene. This indicates that the endogenous gene loci of the animal has been modified and substantially all cells contain such modification.

 Functional disruption of the gene can be accomplished in any effective way,
20 including, e.g., introduction of a stop codon into any part of the coding sequence such that the resulting polypeptide is biologically inactive (e.g., because it lacks a catalytic domain, a ligand binding domain, etc.), introduction of a mutation into a promoter or other regulatory sequence that is effective to turn it off, or reduce transcription of the gene, insertion of an exogenous sequence into the gene which inactivates it (e.g., which disrupts the production of
25 a biologically-active polypeptide or which disrupts the promoter or other transcriptional machinery), deletion of sequences from the gene (or homolog thereof), etc. Examples of transgenic animals having functionally disrupted genes are well known, e.g., as described in U.S. Pat. Nos. 6,239,326, 6,225,525, 6,207,878, 6,194,633, 6,187,992, 6,180,849, 6,177,610, 6,100,445, 6,087,555, 6,080,910, 6,069,297, 6,060,642, 6,028,244, 6,013,858, 5,981,830,
30 5,866,760, 5,859,314, 5,850,004, 5,817,912, 5,789,654, 5,777,195, and 5,569,824. A transgenic animal which comprises the functional disruption can also be referred to as a

-67-

“knock-out” animal, since the biological activity has been “knocked-out.” Knock-outs can be homozygous or heterozygous.

For creating functionally disrupted genes, and other gene mutations, homologous recombination technology is of special interest since it allows specific regions of the genome to be targeted. Using homologous recombination methods, genes can be specifically-inactivated, specific mutations can be introduced, and exogenous sequences can be introduced at specific sites. These methods are well known in the art, e.g., as described in the patents above. See, also, Robertson, *Biol. Reproduc.*, 44(2):238-245, 1991. Generally, the genetic engineering is performed in an embryonic stem (ES) cell, or other pluripotent cell line (e.g., adult stem cells, EG cells), and that genetically-modified cell (or nucleus) is used to create a whole organism. Nuclear transfer can be used in combination with homologous recombination technologies.

For example, the polynucleotides locus can be disrupted in mouse ES cells using a positive-negative selection method (e.g., Mansour et al., *Nature*, 336:348-352, 1988). In this method, a targeting vector can be constructed which comprises a part of the gene to be targeted. A selectable marker, such as neomycin resistance genes, can be inserted into a polynucleotides exon present in the targeting vector, disrupting it. When the vector recombines with the ES cell genome, it disrupts the function of the gene. The presence in the cell of the vector can be determined by expression of neomycin resistance. See, e.g., U.S. Pat. No. 6,239,326. Cells having at least one functionally disrupted gene can be used to make chimeric and germline animals, e.g., animals having somatic and/or germ cells comprising the engineered gene. Homozygous knock-out animals can be obtained from breeding heterozygous knock-out animals. See, e.g., U.S. Pat. No. 6,225,525.

The present invention also relates to non-human, transgenic animal whose genome comprises recombinant polynucleotides nucleic acid (and homologs thereof) operatively linked to an expression control sequence effective to express said coding sequence. Such a transgenic animal can also be referred to as a “knock-in” animal since an exogenous gene has been introduced, stably, into its genome.

A recombinant nucleic acid refers to a polynucleotide which has been introduced into a target host cell and optionally modified, such as cells derived from animals, plants, bacteria, yeast, etc. A recombinant nucleic acid includes completely synthetic nucleic acid sequences,

-68-

semi-synthetic nucleic acid sequences, sequences derived from natural sources, and chimeras thereof. "Operable linkage" has the meaning used through the specification, i.e., placed in a functional relationship with another nucleic acid. When a gene is operably linked to an expression control sequence, as explained above, it indicates that the gene (e.g., coding sequence) is joined to the expression control sequence (e.g., promoter) in such a way that facilitates transcription and translation of the coding sequence. As described above, the phrase "genome" indicates that the genome of the cell has been modified. In this case, the recombinant polynucleotide has been stably integrated into the genome of the animal. The nucleic acid (e.g., a coding sequence) in operable linkage with the expression control sequence can also be referred to as a construct or transgene.

Any expression control sequence can be used depending on the purpose. For instance, if selective expression is desired, then expression control sequences which limit its expression can be selected. These include, e.g., tissue or cell-specific promoters, introns, enhancers, etc. For various methods of cell and tissue-specific expression, see, e.g., U.S. Pat. Nos. 6,215,040, 6,210,736, and 6,153,427. These also include the endogenous promoter, i.e., the coding sequence can be operably linked to its own promoter. Inducible and regulatable promoters can also be utilized.

The present invention also relates to a transgenic animal which contains a functionally disrupted and a transgene stably integrated into the animal's genome. Such an animal can be constructed using combinations any of the above- and below-mentioned methods. Such animals have any of the aforementioned uses, including permitting the knock-out of the normal gene and its replacement with a mutated gene. Such a transgene can be integrated at the endogenous gene locus so that the functional disruption and "knock-in" are carried out in the same step.

In addition to the methods mentioned above, transgenic animals can be prepared according to known methods, including, e.g., by pronuclear injection of recombinant genes into pronuclei of 1-cell embryos, incorporating an artificial yeast chromosome into embryonic stem cells, gene targeting methods, embryonic stem cell methodology, cloning methods, nuclear transfer methods. See, also, e.g., U.S. Patent Nos. 4,736,866; 4,873,191; 4,873,316; 5,082,779; 5,304,489; 5,174,986; 5,175,384; 5,175,385; 5,221,778; Gordon et al., Proc. Natl. Acad. Sci., 77:7380-7384, 1980; Palmiter et al., Cell, 41:343-345, 1985; Palmiter

et al., *Ann. Rev. Genet.*, 20:465-499, 1986; Askew et al., *Mol. Cell. Bio.*, 13:4115-4124, 1993; Games et al. *Nature*, 373:523-527, 1995; Valancius and Smithies, *Mol. Cell. Bio.*, 11:1402-1408, 1991; Stacey et al., *Mol. Cell. Bio.*, 14:1009-1016, 1994; Hasty et al., *Nature*, 350:243-246, 1995; Rubinstein et al., *Nucl. Acid Res.*, 21:2613-2617, 1993; Cibelli et al.,
5 *Science*, 280:1256-1258, 1998. For guidance on recombinase excision systems, see, e.g., U.S. Pat. Nos. 5,626,159, 5,527,695, and 5,434,066. See also, Orban, P.C., et al., "Tissue- and Site-Specific DNA Recombination in Transgenic Mice," *Proc. Natl. Acad. Sci. USA*, 89:6861-6865 (1992); O'Gorman, S., et al., "Recombinase-Mediated Gene Activation and Site-Specific Integration in Mammalian Cells," *Science*, 251:1351-1355 (1991); Sauer, B., et
10 al., "Cre-stimulated recombination at loxP-Containing DNA sequences placed into the mammalian genome," *Polynucleotides Research*, 17(1):147-161 (1989); Gagnetten, S. et al. (1997) *Nucl. Acids Res.* 25:3326-3331; Xiao and Weaver (1997) *Nucl. Acids Res.* 25:2985-2991; Agah, R. et al. (1997) *J. Clin. Invest.* 100:169-179; Barlow, C. et al. (1997) *Nucl. Acids Res.* 25:2543-2545; Araki, K. et al. (1997) *Nucl. Acids Res.* 25:868-872; Mortensen,
15 R. N. et al. (1992) *Mol. Cell. Biol.* 12:2391-2395 (G418 escalation method); Lakhiani, P. P. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:9950-9955 ("hit and run"); Westphal and Leder (1997) *Curr. Biol.* 7:530-533 (transposon-generated "knock-out" and "knock-in"); Templeton, N. S. et al. (1997) *Gene Ther.* 4:700-709 (methods for efficient gene targeting, allowing for a high frequency of homologous recombination events, e.g., without selectable
20 markers); PCT International Publication WO 93/22443 (functionally-disrupted).

A polynucleotide according to the present invention can be introduced into any non-human animal, including a non-human mammal, mouse (Hogan et al., Manipulating the Mouse Embryo: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1986), pig (Hammer et al., *Nature*, 315:343-345, 1985), sheep (Hammer et al.,
25 *Nature*, 315:343-345, 1985), cattle, rat, or primate. See also, e.g., Church, 1987, *Trends in Biotech.* 5:13-19; Clark et al., *Trends in Biotech.* 5:20-24, 1987); and DePamphilis et al., *BioTechniques*, 6:662-680, 1988. Transgenic animals can be produced by the methods described in U.S. Pat. No. 5,994,618, and utilized for any of the utilities described therein.

30 Database

The present invention also relates to electronic forms of polynucleotides,

polypeptides, etc., of the present invention, including computer-readable medium (e.g., magnetic, optical, etc., stored in any suitable format, such as flat files or hierarchical files) which comprise such sequences, or fragments thereof, e-commerce-related means, etc. Along these lines, the present invention relates to methods of retrieving gene sequences from a computer-readable medium, comprising, one or more of the following steps in any effective order, e.g., selecting a cell or gene expression profile, e.g., a profile that specifies that said gene is expressed in a particular (see the expression profiles described above), and retrieving said differentially expressed gene sequences, where the gene sequences consist of the genes represented by SEQ ID NOS 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, and/or 58, or the polypeptides encoded thereby.

A "gene expression profile" means the list of tissues, cells, etc., in which a defined gene is expressed (i.e, transcribed and/or translated). A "cell expression profile" means the genes which are expressed in the particular cell type. The profile can be a list of the tissues in which the gene is expressed, but can include additional information as well, including level of expression (e.g., a quantity as compared or normalized to a control gene), and information on temporal (e.g., at what point in the cell-cycle or developmental program) and spatial expression. By the phrase "selecting a gene or cell expression profile," it is meant that a user decides what type of gene or cell expression pattern he is interested in retrieving. Any pattern of expression preferences may be selected. The selecting can be performed by any effective method. In general, "selecting" refers to the process in which a user forms a query that is used to search a database of gene expression profiles. The step of retrieving involves searching for results in a database that correspond to the query set forth in the selecting step. Any suitable algorithm can be utilized to perform the search query, including algorithms that look for matches, or that perform optimization between query and data. The database is information that has been stored in an appropriate storage medium, having a suitable computer-readable format. Once results are retrieved, they can be displayed in any suitable format, such as HTML. A query is formed by the user to retrieve the set of genes from the database having the desired property. Once the query is inputted into the system, a search algorithm is used to interrogate the database, and retrieve results.

30

Advertising, licensing, etc., methods

-71-

The present invention also relates to methods of advertising, licensing, selling, purchasing, brokering, etc., genes, polynucleotides, specific-binding partners, antibodies, etc., of the present invention. Methods can comprises, e.g., displaying a gene, polynucleotide, polypeptide, or antibody specific for a polypeptide in a printed or computer-readable medium (e.g., on the Web or Internet), accepting an offer to purchase said gene, polypeptide, or antibody, etc.

Other

A polynucleotide, probe, polypeptide, antibody, specific-binding partner, etc., according to the present invention can be isolated. The term "isolated" means that the material is in a form in which it is not found in its original environment or in nature, e.g., more concentrated, more purified, separated from component, etc. An isolated polynucleotide includes, e.g., a polynucleotide having the sequenced separated from the chromosomal DNA found in a living animal, e.g., as the complete gene, a transcript, or a cDNA. This polynucleotide can be part of a vector or inserted into a chromosome (by specific gene-targeting or by random integration at a position other than its normal position) and still be isolated in that it is not in a form that is found in its natural environment. A polynucleotide, polypeptide, etc., of the present invention can also be substantially purified. By substantially purified, it is meant that polynucleotide or polypeptide is separated and is essentially free from other polynucleotides or polypeptides, i.e., the polynucleotide or polypeptide is the primary and active constituent. A polynucleotide can also be a recombinant molecule. By "recombinant," it is meant that the polynucleotide is an arrangement or form which does not occur in nature. For instance, a recombinant molecule comprising a promoter sequence would not encompass the naturally-occurring gene, but would include the promoter operably linked to a coding sequence not associated with it in nature, e.g., a reporter gene, or a truncation of the normal coding sequence.

The term "marker" is used herein to indicate a means for detecting or labeling a target. A marker can be a polynucleotide (usually referred to as a "probe"), polypeptide (e.g., an antibody conjugated to a detectable label), PNA, or any effective material.

-72-

The topic headings set forth above are meant as guidance where certain information can be found in the application, but are not intended to be the only source in the application where information on such topic can be found. Reference materials

For other aspects of the polynucleotides, reference is made to standard textbooks of
5 molecular biology. See, e.g., Hames et al., Polynucleotide Hybridization, IL Press, 1985; Davis et al., Basic Methods in Molecular Biology, Elsevier Sciences Publishing, Inc., New York, 1986; Sambrook et al., Molecular Cloning, CSH Press, 1989; Howe, Gene Cloning and Manipulation, Cambridge University Press, 1995; Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, Inc., 1994-1998.

10 Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever. The entire disclosure of
15 all applications, patents and publications, cited above and in the figures are hereby incorporated by reference in their entirety, including U.S. Serial No. 10/094,092, filed March 11, 2002, U.S. Serial No. 10/112,372, filed April 1, 2002, U.S. Serial No. 60/382, 614, filed May 24, 2002, U.S. Serial No. 10/164,717, filed June 10, 2002, U.S. Serial No. 10/167,631, filed June 13, 2002, U.S. Serial No. 10/177,917, filed June 24, 2002, and U.S. Serial No. 60/399,125, filed 30 July 2002, which are hereby incorporated by reference in their entirety.

20

TABLE 1

Variant ^a	Nucleotide change
Pro18Ala	52C > G
86insA	
Val60Leu	178G > T
Ala64Ser	190G > T
Arg67Gln	200G > A
Phe76Tyr	227T > A
Asp84Glu	252C > A
Ala81Pro	241G > C
Val92Met	274G > A
Thr95Met	284C > T
Val97Ile	289G > A
Ala103Val	308C > T
Gly104Ser	310G > A
Leu106Gln	317T > A
Leu106Leu	318G > A
Arg142His	425G > A
Arg151Cys	451C > T
Arg151Arg	453C > G
Ile155Thr	464T > C
Arg160Trp	478C > T
Arg163Gln	488G > A
Val173del	
Val174Ile	520G > A
537insC	
Pro230Leu	689C > T
Pro230Pro	690G > A
Gln233Gln	699G > A
His260Pro	779A > C
Ile264Ile	792C > T
Cys273Cys	819C > T
Lys278Glu	832A > G
Asn279Ser	836A > G
Asn279Lys	837C > A
Ile287Met	861C > G
Asp294His	880G > C
Phe300Phe	900C > T
Thr314Thr	942A > G
Set316Ser	948C > T

TABLE 2

Allele	Allele Frequency, %		Stimulation of cAMP Production
	White Populations	Individuals With Red Hair	
Wild type	53	23	+++
Val60Leu†	10	3	+
Ala64Ser	<1	1	NA
Lys65Asn	<1	<1	NA
Arg67Gln	0‡	0	NA
Arg67Val	0‡	0	NA
Phe76Tyr	<1	<1	NA
Asp84Glu	1	3	+++
Asn91Asp	<1	0	NA
Val92Leu	<1	1	NA
Val92Met	8	8	+++
Thr95Met	<1	1	NA
Val97Ile	<1	<1	NA
Ala103Val	<1	<1	NA
Leu106Gln	<1	<1	NA
Arg142His	<1	1	-
Arg151Cys§	8	25	-
Ile155Thr	<1	<1	NA
Arg160Trp§	7	19	-
Arg163Gln	4	<1	NA
Ile287Met	0‡	0	NA
Asp294His§	4	13	-
Ala299Thr	<1	<1	NA
ins29	<1	<1	-
ins179	<1	<1	-

*Several synonymous variants have also been described, including Leu106Leu, Leu158Leu, Gln233Gln, Cys273Cys, Phe300Phe, Thr314Thr, and Ser316Ser. MC1-R indicates melanocortin-1 receptor; cAMP, cyclic adenosine monophosphate; triple plus sign, significant stimulation (same as wild type); single plus sign, minimal stimulation; NA, data not available; and minus sign, no stimulation (nonfunctional receptor).

†Possible association with blond/fair hair.

‡Present in <1% of East/Southeast Asians.

§Strong association with red hair, fair skin, and poor tanning ability; recent work also shows an association with cutaneous melanoma and nonmelanoma skin cancer.

||Present in >70% of East/Southeast Asian and Native Americans.

||ins indicates insertion; these single-nucleotide insertion mutations produce frameshifts that result in a prematurely terminated, nonfunctioning

Claims:

1. An isolated polynucleotide comprising,
a polynucleotide sequence which codes without interruption for an amino acid
sequence set forth in SEQ ID NO 2, 12, 17, 26, 39, 44, 46, 48, 50, 52, or 59, or a complement
5 thereto.
2. An isolated polynucleotide of claim 1, comprising a polynucleotide sequence set forth in
SEQ ID NO 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, or 58, or a complement thereto.
- 10 3. An isolated polynucleotide comprising,
a polynucleotide sequence having 95% or more sequence identity along the entire
length of the polynucleotide sequence set forth in SEQ ID NO 1, 11, 16, 25, 38, 43, 45, 47,
49, 51, or 58 of claim 1, or a complement thereto.
- 15 4. An isolated polynucleotide comprising,
a human polynucleotide sequence which hybridizes under high stringency conditions
to a polynucleotide having a polynucleotide sequence set forth in SEQ ID NO 1, 11, 16, 25,
38, 43, 45, 47, 49, 51, or 58 of claim 1, or a complement thereto.
- 20 5. An isolated polynucleotide of claim 4, wherein said high stringency conditions comprise
hybridizing 42°C in 5X SSPE, 0.3% SDS, and 50% formamide, and washes at 65°C for 15
minutes in 2X SSC, and 0.2% SDS.
6. An isolated polypeptide comprising,
25 the amino acid sequence set forth in SEQ ID NOS 2, 12, 17, 26, 39, 44, 46, 48, 50,
52, or 59.
7. An isolated polypeptide comprising,
an amino acid sequence having 95% or more sequence identity along the entire length
30 of the amino acid sequence of claim 6.

8. An isolated polypeptide which is coded for by a polynucleotide of claim 4.

9. A method of detecting a nucleic acid coding, comprising,

5 contacting a sample comprising nucleic acid with a polynucleotide probe specific for
a human muscle selective polynucleotide of claim 1 under conditions effective for said probe
to hybridize specifically with said polynucleotide, and
 detecting hybridization between said probe and said nucleic acid.

10. A method of claim 9, wherein said detecting is performed by:

10 Northern blot analysis, polymerase chain reaction (PCR), reverse transcriptase PCR,
RACE PCR, or *in situ* hybridization.

11. A method of diagnosing a disease associated with abnormal expression of a gene in a
subject, or determining a subject's susceptibility to such disease, comprising:

15 assessing the expression of said gene in said subject.

12. A method of claim 11, wherein assessing is:

 measuring expression levels of said gene, determining the genomic structure of said
gene, determining the mRNA structure of transcripts from said gene, or measuring the
20 expression levels of polypeptide coded for by said gene, and

13. A method of claim 11, wherein said assessing is performed by:

 Northern blot analysis, polymerase chain reaction (PCR), reverse transcriptase PCR,
RACE PCR, or *in situ* hybridization, and
25 using a polynucleotide probe having a polynucleotide sequence selected from SEQ ID
NO 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, or 58, or a complement thereto.

14. A method for identifying an agent that modulates the expression of a gene in a cell,
comprising,

30 contacting a cell population with a test agent under conditions effective for said test
agent to modulate the expression of a polynucleotide of claim 1 in said cells, and

determining whether said test agent modulates said polynucleotide.

15. A method for identifying an agent that modulates the expression of a polypeptide coded for a gene, comprising,

5 contacting a polypeptide coded for by a polynucleotide of claim 1, with a test agent under conditions effective for said test agent to modulate said polypeptide, and
 determining whether said test agent modulates said polypeptide.

16. A method of claim 15, wherein said test agent is an antibody.

10

17. A method of identifying a genetic basis for a disease or disease-susceptibility, comprising:

 determining the association of a disease or disease-susceptibility with a polynucleotide of claim 1.

15

18. A method of claim 17, wherein determining is performed by producing a human-linkage map using said polynucleotide.

20

19. A method of claim 17, wherein determining is performed by comparing the nucleotide sequences of said polynucleotide between normal subjects and subjects having a muscle disease.

25

20. A non-human, transgenic mammal, or a cell thereof, whose genome comprises a functional disruption of a homolog of a gene of claim 1.

30

21. A method of advertising genes for sale, commercial use, or licensing, comprising,
 displaying in a computer-readable medium a polynucleotide set forth in SEQ ID NO 1, 11, 16, 25, 38, 43, 45, 47, 49, 51, or 58 of claim 1, or complements thereto.

22. A method of selecting a polynucleotide sequence coding for a polypeptide, or a polypeptide sequence thereof, from a database comprising polynucleotide sequences and/or

polypeptide sequences, comprising

displaying, in a computer-readable medium, a polynucleotide sequence of claim 1, or polypeptide encoded thereby, or complements to the polynucleotides sequence,

wherein said displayed sequences have been retrieved from said database upon
5 selection by a user.

23. An antibody which is specific for a polypeptide of claim 6, 7, or 8.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

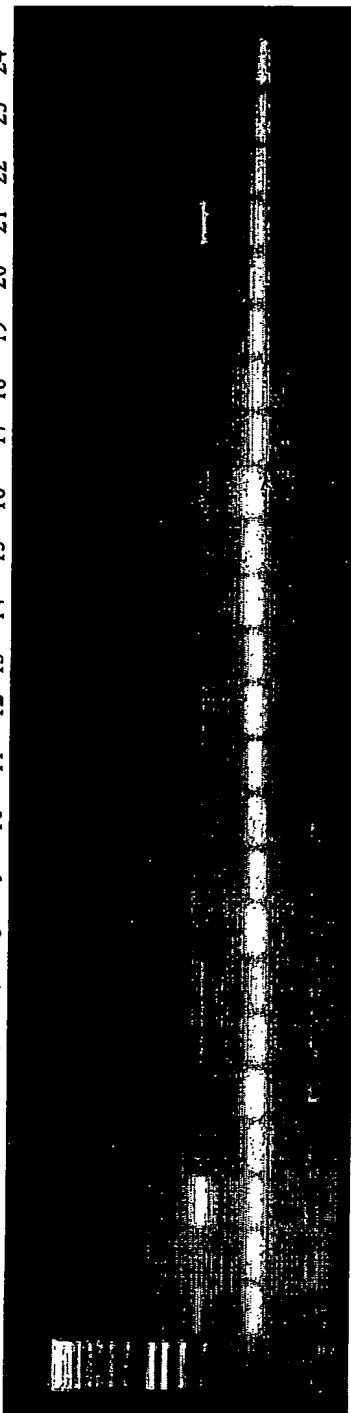


FIG 1

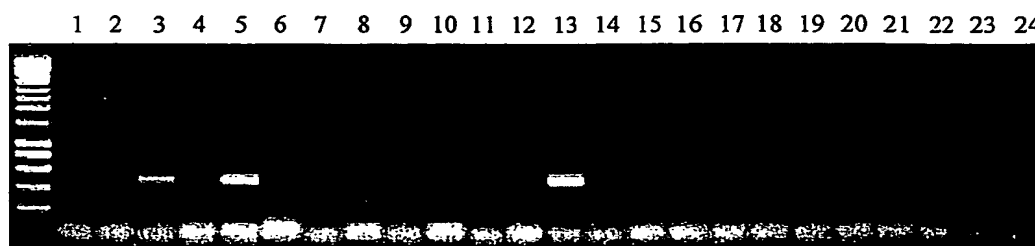


FIG 2

```

*      20      *      40      *      60
OTB182 : MGSTMEPPGGAYLHLGAVTSPVGTARVLQLAFGCTTESLVHRGGFAGVQGTFECMAANGFC : 61
AK003645 : MGSTMEPPGGAYLHLGAVTSPVGTARVLQLAFGCTTESLVHRGGFAGVQGTFECMAANGFC : 61

*      80      *      100      *      120
OTB182 : FAVSALVVACEETRLHGCLRLSWGNETAAEAMLATLLCATAAATYPLYFARRECPPEPAGC : 122
AK003645 : FAFSVLVVACEETKLHSCRLSWGNETAAEAMLATLLCATAAATYPLYFTRLECPPEPAGC : 122

*      140      *      160      *      180
OTB182 : AARDERLAASVFAGLLELAYAVEVALTRARPGQVSSYMATVSGLLKIVQAFVACIIFGALV : 183
AK003645 : MVAPCQRPAPESPWKDDDDVMTAMEYLSRHPT----- : 153

*      200      *      220      *      240
OTB182 : HDSRYGRYVATQMCVAVYSLCFLATVAVVALSVMGHTGGLGCPEDRLVWVYTFEVLALLYLS : 244
AK003645 : ----- :

*      260      *      280      *      300
OTB182 : AAVIMPVECFDPKYGEPKRPPNCARGSCPWDSQLVVAIFTYVNLVWDLAYSQRIFVP : 305
AK003645 : ----- :

OTB182 : SL : 307
AK003645 : -- : -

```

FIG 3

		*	20	*	40	*	6	
TRPCC	:	MPEPWGT	VYFLGIAQVFS	FLFSWNN	LEGVMNQADAPRPLN	WTIRKLCHAAFLPSVRL	LLK	: 59
AB046836	:	-----	-----	-----	-----	-----	-----	: -
XM_036123	:	-----	-----	-----	-----	-----	-----	: -
XM_140575	:	-----	-----	-----	-----	-----	-----	: -
		0	*	80	*	100	*	1
TRPCC	:	AQKSWIERAFYKRECVHIIPSTKDPHRC	CCGRLIGQHVGLT	PSISVLQNEKNESRLSRN	:			: 118
AB046836	:	-----	-----	-----	-----	-----	-----	: -
XM_036123	:	-----	-----	-----	-----	-----	-----	: -
XM_140575	:	-----	-----	-----	-----	-----	-----	: -
		20	*	140	*	160	*	
TRPCC	:	DIQSEKWSISKHTQLSP	TDAGTIEFQGGGHSNKAMYVRVSE	DTKPDLLLHLM	TKEWQL	:		: 177
AB046836	:	-----	-----	-----	-----	-----	-----	: -
XM_036123	:	-----	-----	-----	-----	-----	-----	: -
XM_140575	:	-----	-----	-----	-----	-----	-----	: -
		180	*	200	*	220	*	
TRPCC	:	ELPKLLISVHGG	LONFELQPKLKQVFGKGLIKAAMTTGAWIFTGGVNTGVIRHVG	DALK	:			: 236
AB046836	:	-----	-----	-----	-----	-----	-----	: -
XM_036123	:	-----	-----	-----	-----	-----	-----	: -
XM_140575	:	-----	-----	-----	-----	-----	-----	: -
		240	*	260	*	280	*	
TRPCC	:	DHASKSRGKICTIGIAPW	GIVENQEDLIGRDVVRPYQTMSNPMSKLTVLNSMHSHEILA	:				: 295
AB046836	:	-----	-----	-----	-----	-----	-----	: -
XM_036123	:	-----	-----	-----	-----	-----	-----	: -
XM_140575	:	-----	-----	-----	-----	-----	-----	: -
		300	*	320	*	340	*	
TRPCC	:	DNGTTGKYGA	EVKLRRQLEKHISLQKINTRIGQGV	PVVALIVEGGPNVISIVLEYLRDT	:			: 354
AB046836	:	-----	-----	-----	-----	-----	-----	: -
XM_036123	:	-----	-----	-----	-----	-----	-----	: -
XM_140575	:	-----	-----	-----	-----	-----	-----	: -
		360	*	380	*	400	*	
TRPCC	:	PPVPVVVCDGSGRASDILAFGHKYSEEGGLINESLRDQLLV	TIQKTFETYTRTQAQHLFI	:				: 413
AB046836	:	-----	-----	-----	-----	-----	-----	: -
XM_036123	:	-----	-----	-----	-----	-----	-----	: -
XM_140575	:	-----	-----	-----	-----	-----	-----	: -
		420	*	440	*	460	*	
TRPCC	:	ILMECMKKKELITVFRMGSEGHQDIDLAILTALLKGANASAPDQLSLALAWN	NRVDIARS	:				: 472
AB046836	:	-----	-----	-----	-----	-----	-----	: -
XM_036123	:	-----	-----	-----	-----	-----	-----	: -
XM_140575	:	-----	-----	-----	-----	-----	-----	: -

FIG. 4A

	480	*	500	*	520	*		
TRPCC	:	QIFIYQQQW	PVGSLEQ	AMLDALV	LDVDFV	KLLIENG	VSMHRFLTISRLEELYNTRHG	531
AB046836	:							-
XM_036123	:							-
XM_140575	:							-

	540	*	560	*	580	*		
TRPCC	:	SNTLYHLV	RDVKKGNL	PPDYRIS	LIDIGLV	EYLMGG	AYRCNYTRKRFR	590
AB046836	:							-
XM_036123	:							-
XM_140575	:							-

	600	*	620	*	640			
TRPCC	:	RPKALKLL	GMEDDI	PLRRGR	KTTKKR	EEVDIDL	DDPEINHF	649
AB046836	:							-
XM_036123	:							-
XM_140575	:							-

	*	660	*	680	*	700		
TRPCC	:	KMALFFW	QHGE	EAMAKAL	VACKLCK	AMAHEAS	ENDMVDDISQ	708
AB046836	:							18
XM_036123	:							-
XM_140575	:							-

	*	720	*	740	*	760		
TRPCC	:	DQSYKQ	DEQLA					767
AB046836	:	DQSYKQ	DEQLA					77
XM_036123	:							48
XM_140575	:							-

	*	780	*	800	*	820		
TRPCC	:							826
AB046836	:							136
XM_036123	:							107
XM_140575	:							-

	*	840	*	860	*	880		
TRPCC	:							885
AB046836	:							195
XM_036123	:							166
XM_140575	:							-

FIG. 4B

		*	900	*	920	*	940	
TRPCC	:							: 944
AB046836	:							: 254
XM_036123	:							: 225
XM_140575	:							: -
		*	960	*	980	*	1000	
TRPCC	:							: 1003
AB046836	:							: 313
XM_036123	:							: 284
XM_140575	:							: 21
		*	1020	*	1040	*	1060	
TRPCC	:							: 1062
AB046836	:							: 372
XM_036123	:							: 343
XM_140575	:							: 80
		*	1080	*	1100	*	1120	
TRPCC	:							: 1121
AB046836	:							: 431
XM_036123	:							: 402
XM_140575	:							: 139
		*	1140	*	1160	*	1180	
TRPCC	:							: 1180
AB046836	:							: 490
XM_036123	:							: 461
XM_140575	:							: 198
		*	1200	*	1220	*	124	
TRPCC	:							: 1239
AB046836	:							: 549
XM_036123	:							: 520
XM_140575	:							: 257
		0	*	1260	*	1280	*	13
TRPCC	:							: 1298
AB046836	:							: 608
XM_036123	:							: 579
XM_140575	:							: 316

FIG. 4C

	00	*	1320	*	1340	*	1	
TRPCC	:	SSFNSEQGNTFKLQESIDPAGEETMSPTSPTLMPMRSHSFYSVNMKDKGGIEKLESIF						: 1357
AB046836	:	SSFNSEQGNTFKLQESIDPAGEETMSPTSPTLMPMRSHSFYSVNMKDKGGIEKLESIF						: 667
XM_036123	:	SSFNSEQGNTFKLQESIDPAGEETMSPTSPTLMPMRSHSFYSVNMKDKGGIEKLESIF						: 638
XM_140575	:	SSFNSEQGNTFKLQESIDPAGEETISPTSPTLMPMRSHSFYSVNVKDKGGIEKLESIF						: 375

	360	*	1380	*	1400	*	
TRPCC	:	KERSLSLHRATSSHSVAKEPKAPAAPANTLAIVPDSRRPSSCIDIYVSAMDELHCDI					: 1416
AB046836	:	KERSLSLHRATSSHSVAKEPKAPAAPANTLAIVPDSRRPSSCIDIYVSAMDELHCDI					: 726
XM_036123	:	KERSLSLHRATSSHSVAKEPKAPAAPANTLAIVPDSRRPSSCIDIYVSAMDELHCDI					: 697
XM_140575	:	KERSLSLHRATSSHSVAKEPKAPAAPANTLAIVPDSRRPSSCIDIYVSAMDELHCDI					: 434

	1420	*	1440	*	1460	*	
TRPCC	:	LDNSVNILGLGEPSESSTPSSAYATLAPTDRPPSRSIDFEDITSMDTRSFSSD					: 1475
AB046836	:	LDNSVNILGLGEPSESSTPSSAYATLAPTDRPPSRSIDFEDITSMDTRSFSSD					: 785
XM_036123	:	LDNSVNILGLGEPSESSTPSSAYATLAPTDRPPSRSIDFEDITSMDTRSFSSD					: 756
XM_140575	:	LDNSMNILGLGEPSESALAPSTPSSAYATLAPTDRPPSRSIDFEDLTSMMDTRSFSSD					: 493

	1480	*	1500	*	1520	*	
TRPCC	:	YTHLPECQNPWDSPEPMYHTIERSKSSRYLATTFELLEEAPIVKSHSFNFSPPSRSYAN					: 1534
AB046836	:	YTHLPECQNPWDSPEPMYHTIERSKSSRYLATTFELLEEAPIVKSHSFNFSPPSRSYAN					: 844
XM_036123	:	YTHLPECQNPWDSPEPMYHTIERSKSSRYLATTFELLEEAPIVKSHSFNFSPPSRSYAN					: 815
XM_140575	:	YTHLPECQNPWDTDPETMYHTIERSKSSRYLATTFELLEEAPIVKSHSFNFSPPSRSYAN					: 552

	1540	*	1560	*	1580	*	
TRPCC	:	FGVPVKTAEYTSITDCIDTRCVNAPQAIADRAEPGGLGDKVEDLTCCHPEREAELSHP					: 1593
AB046836	:	FGVPVKTAEYTSITDCIDTRCVNAPQAIADRAEPGGLGDKVEDLTCCHPEREAELSHP					: 903
XM_036123	:	FGVPVKTAEYTSITDCIDTRCVNAPQAIADRAEPGGLGDKVEDLTCCHPEREAELSHP					: 874
XM_140575	:	FGVPVKTAEYTSITDCIDTRCVNAPQAIADRAEPGGLGDKVEDLSCCHPEREAELSHP					: 611

	1600	*	1620	*	1640	*	
TRPCC	:	SSDSEENEAKGRANFISSQEAENRTLSNNITVPKIERANSYSAEENVPYAHTRK					: 1652
AB046836	:	SSDSEENEAKGRANFISSQEAENRTLSNNITVPKIERANSYSAEENVPYAHTRK					: 962
XM_036123	:	SSDSEENEAKGRANFISSQEAENRTLSNNITVPKIERANSYSAEENVPYAHTRK					: 933
XM_140575	:	SSDSEENEARGORANFISSQEAENADRTLSNNITVPKIERANSYSAEENVPYAHTRK					: 670

	1660	*	1680	*	1700	*	
TRPCC	:	SFSISDKLDRQRNTASLRNPFQRTKILQYTPNKLYPECLLSSTGAVELMDPAEAILLA					: 1707
AB046836	:	SFSISDKLDRQRNTASLRNPFQRTKILQYTPNKLYPECLLSSTGAVELMDPAEAILLA					: 1017
XM_036123	:	SFSISDKLDRQRNTASLRNPFQRTKILQYTPNKLYPECLLSSTGAVELMDPAEAILLA					: 988
XM_140575	:	SFSISDKLDRQRNTASLRNPFQRTKILQYTPNKLYPECLLSSTGAVELMDPAEAILLA					: 729

	1720	*	
TRPCC	:	-----	: -
AB046836	:	-----	: -
XM_036123	:	-----	: -
XM_140575	:	AFLDGGY-----	: 736

FIG. 4D



FIG. 5

```

      *           20           *           40           *           60
NM_002386 : MAVQGSQRRLLGSLNSTPTAI PQLGLAANQTGARCLEVSI SDGLFLSLGLVSLVENALVV : 60
MC-1RC    : MAVQGSQRRLLGSLNSTPTAI PQLGLAANQTGARCLEVSI SDGLFLSLGLVSLVENALVV : 60
MC-1RB    : MAVQGSQRRLLGSLNSTPTAI PQLGLAANQTGARCLEVSI SDGLFLSLGLVSLVENALVV : 60

      *           80           *           100          *           120
NM_002386 : ATIAKNRNLHSPMYCFICCLALSDLLVSGSNVLETAVILLLEAGALVARAAVLQQLDNVT : 120
MC-1RC    : ATIAKNRNLHSPMYCFICCLALSDLLVSGSNVLETAVILLLEAGALVARAAVLQQLDNVT : 120
MC-1RB    : ATIAKNRNLHSPMYCFICCLALSDLLVSGSNVLETAVILLLEAGALVARAAVLQQLDNVT : 120

      *           140          *           160          *           180
NM_002386 : DVITCSSMLSSLCFLGAI AVDRYISIFYALRYHSIVTLPRARQAVAAIWVASVVFSTLEI : 180
MC-1RC    : DVITCSSMLSSLCFLGAI AVDRYISIFYALRYHSIVTLPRARQAVAAIWVASVVFSTLEI : 180
MC-1RB    : DVITCSSMLSSLCFLGAI AVDRYISIFYALRYHSIVTLPRARQAVAAIWVASVVFSTLEI : 180

      *           200          *           220          *           240
NM_002386 : AYYDHVAVLLCLVVFFLAMLVLMVAVLYVHMLARACQHAQGIARLHKRQRPVHQGFGLKGA : 240
MC-1RC    : AYYDHVAVLLCLVVFFLAMLVLMVAVLYVHMLARACQHAQGIARLHKRQRPVHQGFGLKGA : 240
MC-1RB    : AYYDHVAVLLCLVVFFLAMLVLMVAVLYVHMLARACQHAQGIARLHKRQRPVHQGFGLKGA : 240

      *           260          *           280          *           300
NM_002386 : VTLTILLGIFFLCWGPFELH LTLIVLCPEHPTCGCIFKNENLFLALIIICNAIIDPLIYAF : 300
MC-1RC    : VTLTILLGIFFLCWGPFELH LTLIVLCPEHPTCGCIFKNENLFLALIIICNAIIDPLIYAF : 300
MC-1RB    : VTLTILLGIFFLCWGPFELH LTLIVLCPEHPTCGCIFKNENLFLALIIICNAIIDPLIYAF : 300

      *           320          *           340          *           360
NM_002386 : HSQELRRTLKEVLTCSW----- : 317
MC-1RC    : HSQELRRTLKEVLTCSQDRALVSDVKSILGGSVCQSLLPQQRQERKQDQKASSPALO : 360
MC-1RB    : HSQELRRTLKEVLTCSQDRALVSDVKSILGGSVCQSLLPQQRQERKQDQKASSPALO : 360

      *           380          *
NM_002386 : ----- : -
MC-1RC    : RILQKEPPRGRTSRCSRAPVPSTLDAVLAAEEAGSQPSL : 398
MC-1RB    : RILQKEVHSLPQAKGPGLQEPF----- : 382

```

FIG 6

10/509853

WO 03/085095

10/509853

PCT/US03/09921

10/20

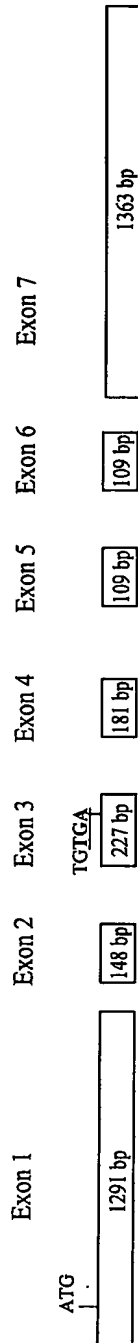


FIG 7

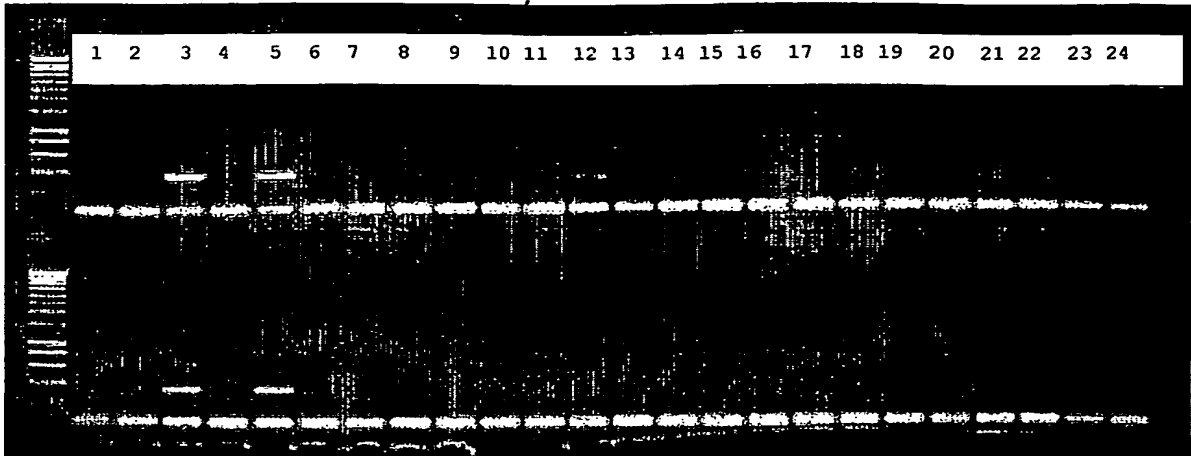


FIG 8

```

      *      20      *      40      *      60
OTB860 : MDGNSLLSVPSNLESSRMYDVLEPQQGRGCGSSGSGPGNSITACKKVLRSNSLLESTDYF : 60
KIAA1678 : ----- : -

      *      80      *      100      *      120
OTB860 : LONQRMPCQIGFVEDKSENCASVCFVNLDVNKDECBTEHLQOKLVNVSFDLEKLISSMNV : 120
KIAA1678 : ----- : -

      *      140      *      160      *      180
OTB860 : QQPKENEIVVLSSGLASGNLQADFEVSQCPWLPDICLVQCARGNRPNSTNCIIFEINKFLI : 180
KIAA1678 : ----- : -

      *      200      *      220      *      240
OTB860 : GLELVQERQLHLETNILKLEDDTNCSSIEEDFLTASEHLEEESEVDESNDYENINVS : 240
KIAA1678 : ----- : -

      *      260      *      280      *      300
OTB860 : ANVLESKQLKGATQVEWNCNKEKWLYALEDKYINKYPTPLIKTERS PENLTKN TALQSLD : 300
KIAA1678 : ----- : -

      *      320      *      340      *      360
OTB860 : PSAKPSQWKREAVGNRQATHYYHSEAFKQOMEKSQALYIPKDAYFSMMDKDVPSACAVA : 360
KIAA1678 : ----- : -

      *      380      *      400      *      420
OTB860 : EQRSNLNPGDHEDTRNALPPRQDGEVTTGKYATNLAESVLQDAFIRLSQSQSTLPQESAV : 420
KIAA1678 : -----DHEDTRNALPPRQDGEVTTGKYATNLAESVLQDAFIRLSQSQSTLPQESAV : 51

      *      440      *      460      *      480
OTB860 : SVSVGSSLLPSCYSTKDTVVSRSWNELPKIVVVQSPDGSDAAPQPGISSWPMEVSVETS : 480
KIAA1678 : SVSVGSSLLPSCYSTKDTVVSRSWNELPKIVVVQSPDGSDAAPQPGISSWPMEVSVETS : 111

      *      500      *      520      *      540
OTB860 : SILSGENSSRQPQSALEVALACAAATVIGTISSPQATERLKMEQVVSNFPPGSSGALQTQA : 540
KIAA1678 : SILSGENSSRQPQSALEVALACAAATVIGTISSPQATERLKMEQVVSNFPPGSSGALQTQA : 171

      *      560      *      580      *      600
OTB860 : PQGLKEPSINEYSFPSALCGMTQVASAVAVCGLGEREEVTCVAPSGSLPPAAEASEAMP : 600
KIAA1678 : PQGLKEPSINEYSFPSALCGMTQVASAVAVCGLGEREEVTCVAPSGSLPPAAEASEAMP : 231

```

FIG. 9A

	*	620	*	640	*	660	
OTB860	:	PLCGLASMELGKEAIAEGLLKEAALVLTRENTYSSIGDFLDSMNRRIMETASKSQTLCSE	:	660			
KIAA1678	:	PLCGLASMELGKEAIAEGLLKEAALVLTRENTYSSIGDFLDSMNRRIMETASKSQTLCSE	:	291			
	*	680	*	700	*	720	
OTB860	:	NVVRNELAHTLSNVILRHSEIDVHHKNMIIDPNDNRHSSEILDTLMESTNQLLLDVICFT	:	720			
KIAA1678	:	NVVRNELAHTLSNVILRHSEIDVHHKNMIIDPNDNRHSSEILDTLMESTNQLLLDVICFT	:	351			
	*	740	*	760	*	780	
OTB860	:	FKKMSHIVRLGECPAVLKETIRRETEPSCQPSDPGASQAWTKATESSSSSPLNSHNT	:	780			
KIAA1678	:	FKKMSHIVRLGECPAVLKETIRRETEPSCQPSDPGASQAWTKATESSSSSPLNSHNT	:	411			
	*	800	*	820	*	840	
OTB860	:	SLVINNLVDGMYSKQDKGGVRPGLFKNPTLQSQLSRSHRVPDSSTATSSKEIYLGKIAG	:	840			
KIAA1678	:	SLVINNLVDGMYSKQDKGGVRPGLFKNPTLQSQLSRSHRVPDSSTATSSKEIYLGKIAG	:	471			
	*	860	*	880	*	900	
OTB860	:	EDTKSPHSENECRASSEGQRSPTVSRSRSGSQEAEESIHPNTQEKYNCATSRINEVQVN	:	900			
KIAA1678	:	EDTKSPHSENECRASSEGQRSPTVSRSRSGSQEAEESIHPNTQEKYNCATSRINEVQVN	:	531			
	*	920	*	940	*	960	
OTB860	:	LSLLGDDLLLPAQSTLQTKHPDIYCITDFAEELADTVVSMATEIAAICLDNSSGKQPWF	:	960			
KIAA1678	:	LSLLGDDLLLPAQSTLQTKHPDIYCITDFAEELADTVVSMATEIAAICLDNSSGKQPWF	:	591			
	*	980	*	1000	*	1020	
OTB860	:	AWKRGSEFLMTPNVPCRSLSKRKKESQSGTAVRKHKPPRLSEIKRKTDEHPDELKEKLMNR	:	1020			
KIAA1678	:	AWKRGSEFLMTPNVPCRSLSKRKKESQSGTAVRKHKPPRLSEIKRKTDEHPDELKEKLMNR	:	651			
	*	1040	*	1060	*	1080	
OTB860	:	VVDESMNLEDVPDSVNLFAVEVAKIMNLTEFSMVDGMWQAQGYPRNRLLSGDRWSRLKA	:	1080			
KIAA1678	:	VVDESMNLEDVPDSVNLFAVEVAKIMNLTEFSMVDGMWQAQGYPRNRLLSGDRWSRLKA	:	711			
	*	1100	*	1120	*	1140	
OTB860	:	SSCESIPEEDSEARAYVNSLGLMSTLSQPVSRASSVSKQSSCESITDEFSRFMVNQMENE	:	1140			
KIAA1678	:	SSCESIPEEDSEARAYVNSLGLMSTLSQPVSRASSVSKQSSCESITDEFSRFMVNQMENE	:	771			
	*	1160	*	1180	*	1200	
OTB860	:	GRGFELLLDYYAGKNASSILNSAMQQACRKS DHLSVRPSCPSKQSSITESITEEFYRYMLR	:	1200			
KIAA1678	:	GRGFELLLDYYAGKNASSILNSAMQQACRKS DHLSVRPSCPSKQSSITESITEEFYRYMLR	:	831			

FIG. 9B

	*	1220	*	1240	*	1260	
OTB860	:	DIERDSRESASSRRSSQDWTAGLLSPSLRSPVCHRQSSMPDSRSPCSRSLTVNVPIKANSL				:	1260
KIAA1678	:	DIERDSRESASSRRSSQDWTAGLLSPSLRSPVCHRQSSMPDSRSPCSRSLTVNVPIKANSL				:	891
	*	1280	*	1300	*	1320	
OTB860	:	DGFAQNC PQDFLSVQPVSSASSSSGLCKSDSCLYRRGGTDHITNMLIHETWASSIEALMRK				:	1320
KIAA1678	:	DGFAQNC PQDFLSVQPVSSASSSSGLCKSDSCLYRRGGTDHITNMLIHETWASSIEALMRK				:	951
	*	1340	*	1360	*	1380	
OTB860	:	NKIIVDDAEADTEPVSGGSPSQAEKCANRLAASRMCSGPTLLVQESLDCPRKDSVTECK				:	1380
KIAA1678	:	NKIIVDDAEADTEPVSGGSPSQAEKCANRLAASRMCSGPTLLVQESLDCPRKDSVTECK				:	1011
	*	1400	*	1420	*	1440	
OTB860	:	QPPVSSL SKTASLTNHSPLDSKKETSSCQDPVPINHKRRSLCSREVPLIQIETDQREACA				:	1440
KIAA1678	:	QPPVSSL SKTASLTNHSPLDSKKETSSCQDPVPINHKRRSLCSREVPLIQIETDQREACA				:	1071
	*	1460	*	1480	*	1500	
OTB860	:	GEPEPFLSKSSLLEEAEGHSNDKNI PDVVRGGDTAVSACQIHSDSLDTRDVPEAEASTEA				:	1500
KIAA1678	:	GEPEPFLSKSSLLEEAEGHSNDKNI PDVVRGGDTAVSACQIHSDSLDTRDVPEAEASTEA				:	1131
	*	1520	*	1540	*	1560	
OTB860	:	RAPDEAPNPPSSSEESTGSWTQLANEEDNPDDTSSFLQLSERSMSGNNGSSATSSSLGIMDI				:	1560
KIAA1678	:	RAPDEAPNPPSSSEESTGSWTQLANEEDNPDDTSSFLQLSERSMS-----				:	1176
	*	1580	*	1600	*	1620	
OTB860	:	DIYQESMPSSPMINELVEEKKILKGQSESTEAPASGPPTGTASPQRSLLVINFDLEPECP				:	1620
KIAA1678	:	-----ELVEEKKILKGQSESTEAPASGPPTGTASPQRSLLVINFDLEPECP				:	1222
	*	1640	*	1660	*	1680	
OTB860	:	DAELRATLQWIAASELGIPTIYFKKSQENRIEKF LDVVQLVHRKSWKVGDI FHAVVQYCK				:	1680
KIAA1678	:	DAELRATLQWIAASELGIPTIYFKKSQENRIEKF LDVVQLVHRKSWKVGDI FHAVVQYCK				:	1282
	*	1700	*	1720			
OTB860	:	MH EEQKDGRLSLFDW LLELG-----				:	1700
KIAA1678	:	MH EEQKDGRLSLFDW LLELG-----				:	1302

FIG 9C

```

      *      20      *      40      *      60
BR137A : MSEQGDLNQAIAEEGGTEQETATPENGIVKSESLEDEEEKLELQRRLEAQNQERRKSKSGAGKGKL : 65
BR137C : MSEQGDLNQAIAEEGGTEQETATPENGIVKSESLEDEEEKLELQRRLEAQNQERRKSKSGAGKGKL : 65
BR137B : MSEQGDLNQAIAEEGGTEQETATPENGIVKSESLEDEEEKLELQRRLEAQNQERRKSKSGAGKGKL : 65
BR137E : MSEQGDLNQAIAEEGGTEQETATPENGIVKSESLEDEEEKLELQRRLEAQNQERRKSKSGAGKGKL : 65
AL133109 : ----- : -
BR137D : MSEQGDLNQAIAEEGGTEQETATPENGIVKSESLEDEEEKLELQRRLEAQNQERRKSKSGAGKGKL : 65

      *      80      *      100      *      120      *
BR137A : TRSLAVCEESSARPGGESLQDQESIHLQLSSFSSSQEEDKSRKDDSEREKEKDKNKDKTSEKPKI : 130
BR137C : TRSLAVCEESSARPGGESLQDQESIHLQLSSFSSSQEEDKSRKDDSEREKEKDKNKDKTSEKPKI : 130
BR137B : TRSLAVCEESSARPGGESLQDQESIHLQLSSFSSSQEEDKSRKDDSEREKEKDKNKDKTSEKPKI : 130
BR137E : TRSLAVCEESSARPGGESLQDQESIHLQLSSFSSSQEEDKSRKDDSEREKEKDKNKDKTSEKPKI : 130
AL133109 : ----- : -
BR137D : TRSLAVCEESSARPGGESLQDQ ----- : 89

      140      *      160      *      180      *
BR137A : RMLSKDCSQEYTDSTGIDLHEFLINTLKNNSRDRMILLKMQEIIDFIADNNNHKKFPQMSSYQ : 195
BR137C : RMLSKDCSQEYTDSTGIDLHEFLINTLKNNSRDRMILLKMQEIIDFIADNNNHKKFPQMSSYQ : 195
BR137B : RMLSKDCSQEYTDSTGIDLHEFLINTLKNNSRDRMILLKMQEIIDFIADNNNHKKFPQMSSYQ : 195
BR137E : RMLSKDCSQEYTDSTGIDLHEFLINTLKNNSRDRMILLKMQEIIDFIADNNNHKKFPQMSSYQ : 195
AL133109 : -----RDRMILLKMQEIIDFIADNNNHKKFPQMSSYQ : 34
BR137D : ----- : -

      200      *      220      *      240      *      260
BR137A : RMLVHRVAAYFGLDHNVDQTGKSVIINKTSSTRIPEQRFCEHLKDEKGEESQKRFLKRDNSSID : 260
BR137C : RMLVHRVAAYFGLDHNVDQTGKSVIINKTSSTRIPEQRFCEHLKDEKGEESQKRFLKRDNSSID : 260
BR137B : RMLVHRVAAYFGLDHNVDQTGKSVIINKTSSTRIPEQRFCEHLKDEKGEESQKRFLKRDNSSID : 260
BR137E : RMLVHRVAAYFGLDHNVDQTGKSVIINKTSSTRIPEQRFCEHLKDEKGEESQKRFLKRDNSSID : 260
AL133109 : RMLVHRVAAYFGLDHNVDQTGKSVIINKTSSTRIPEQRFCEHLKDEKGEESQKRFLKRDNSSID : 99
BR137D : ----- : -

      *      280      *      300      *      320
BR137A : KEDNQQRNRMHPFRDDRRSKSIEEREEYQVRERIFAHDVCSQESLFVEN----- : 311
BR137C : KEDNQ-----VCSQESLFVEN----- : 277
BR137B : KEDNQ-----VCSQESLFVENSRLLEDNSNICETY : 291
BR137E : KEDNQQRNRMHPFRDDRRSKSIEEREEYQVRERIFAHDVCSQESLFVENSRLLEDNSNICETY : 325
AL133109 : KEDNQ-----VCSQESLFVEN-----RLLEDNSNICETY : 129
BR137D : ----- : -

      *      340      *      360      *      380      *
BR137A : -----RGNRDGSGR TSGSRQSSSENELKWS DHQRAWSS TDS DSSNRNLKPAMTKTASFGGITVL : 370
BR137C : -----RGNRDGSGR TSGSRQSSSENELKWS DHQRAWSS TDS DSSNRNLKPAMTKTASFGGITVL : 336
BR137B : KKRQLFRGNRDGSGR TSGSRQSSSENELKWS DHQRAWSS TDS DSSNRNLKPAMTKTASFGGITVL : 356
BR137E : KKRQLFRGNRDGSGR TSGSRQSSSENELKWS DHQRAWSS TDS DSSNRNLKPAMTKTASFGGITVL : 390
AL133109 : KKRQLFRGNRDGSGR TSGSRQSSSENELKWS DHQRAWSS TDS DSSNRNLKPAMTKTASFGGITVL : 194
BR137D : ----- : -

      400      *      420      *      440      *
BR137A : TRGDSTSSSTRSTGKLSKAGSESSSSAGSSGSLSRTHPPLQSTPLVSGVAAGSPGCVYPENGIGG : 435
BR137C : TRGDSTSSSTRSTGKLSKAGSESSSSAGSSGSLSRTHPPLQSTPLVSGVAAGSPGCVYPENGIGG : 401
BR137B : TRGDSTSSSTRSTGKLSKAGSESSSSAGSSGSLSRTHPPLQSTPLVSGVAAGSPGCVYPENGIGG : 421
BR137E : TRGDSTSSSTRSTGKLSKAGSESSSSAGSSGSLSRTHPPLQSTPLVSGVAAGSPGCVYPENGIGG : 455
AL133109 : TRGDSTSSSTRSTGKLSKAGSESSSSAGSSGSLSRTHPPLQSTPLVSGVAAGSPGCVYPENGIGG : 259
BR137D : ----- : -

```

FIG 10A

```

      460          *          480          *          500          *          520
BR137A : QVAPSSTSYILLPLEAATGIPPGSILLNPHTGQPFVNPDTGTPAIYNPPTSQQPLRSAMVGQSQQQ : 500
BR137C : QVAPSSTSYILLPLEAATGIPPGSILLNPHTGQPFVNPDTGTPAIYNPPTSQQPLRSAMVGQSQQQ : 466
BR137B : QVAPSSTSYILLPLEAATGIPPGSILLNPHTGQPFVNPDTGTPAIYNPPTSQQPLRSAMVGQSQQQ : 486
BR137E : QVAPSSTSYILLPLEAATGIPPGSILLNPHTGQPFVNPDTGTPAIYNPPTSQQPLRSAMVGQSQQQ : 520
AL133109 : QVAPSSTSYILLPLEAATGIPPGSILLNPHTGQPFVNPDTGTPAIYNPPTSQQPLRSAMVGQSQQQ : 324
BR137D : ----- : -

      *          540          *          560          *          580
BR137A : PPQQQPSQPQQQVQPPQPMAGPLVTQSVQGLQASSQSVQYPAVSFPPQHLLPVSPTOHFFPMRD : 565
BR137C : PPQQQPSQPQQQVQPPQPMAGPLVTQSVQGLQASSQSVQYPAVSFPPQHLLPVSPTOHFFPMRD : 531
BR137B : PPQQQPSQPQQQVQPPQPMAGPLVTQSVQGLQASSQSVQYPAVSFPPQHLLPVSPTOHFFPMRD : 551
BR137E : PPQQQPSQPQQQVQPPQPMAGPLVTQSVQGLQASSQSVQYPAVSFPPQHLLPVSPTOHFFPMRD : 585
AL133109 : PPQQQPSQPQQQVQPPQPMAGPLVTQSVQGLQASSQSVQYPAVSFPPQHLLPVSPTOHFFPMRD : 389
BR137D : ----- : -

      *          600          *          620          *          640          *
BR137A : DVATQFGQMTLSRQSSGETPEPPSGFVYPSSLMPPQPAQQPSYVIASSTGQQLPTGGFSGSGPPISQ : 630
BR137C : DVATQFGQMTLSRQSSGETPEPPSGFVYPSSLMPPQPAQQPSYVIASSTGQQLPTGGFSGSGPPISQ : 596
BR137B : DVATQFGQMTLSRQSSGETPEPPSGFVYPSSLMPPQPAQQPSYVIASSTGQQLPTGGFSGSGPPISQ : 616
BR137E : DVATQFGQMTLSRQSSGETPEPPSGFVYPSSLMPPQPAQQPSYVIASSTGQQLPTGGFSGSGPPISQ : 650
AL133109 : DVATQFGQMTLSRQSSGETPEPPSGFVYPSSLMPPQPAQQPSYVIASSTGQQLPTGGFSGSGPPISQ : 454
BR137D : ----- : -

      660          *          680          *          700          *
BR137A : QVLQPPPSQGFVQQPPPAQMPVYYYPSGQYPTSTTQQYRPMAPVQYNAQRSQQMPQAAQQAGYQ : 695
BR137C : QVLQPPPSQGFVQQPPPAQMPVYYYPSGQYPTSTTQQYRPMAPVQYNAQRSQQMPQAAQQAGYQ : 661
BR137B : QVLQPPPSQGFVQQPPPAQMPVYYYPSGQYPTSTTQQYRPMAPVQYNAQRSQQMPQAAQQAGYQ : 681
BR137E : QVLQPPPSQGFVQQPPPAQMPVYYYPSGQYPTSTTQQYRPMAPVQYNAQRSQQMPQAAQQAGYQ : 715
AL133109 : QVLQPPPSQGFVQQPPPAQMPVYYYPSGQYPTSTTQQYRPMAPVQYNAQRSQQMPQAAQQAGYQ : 519
BR137D : ----- : -

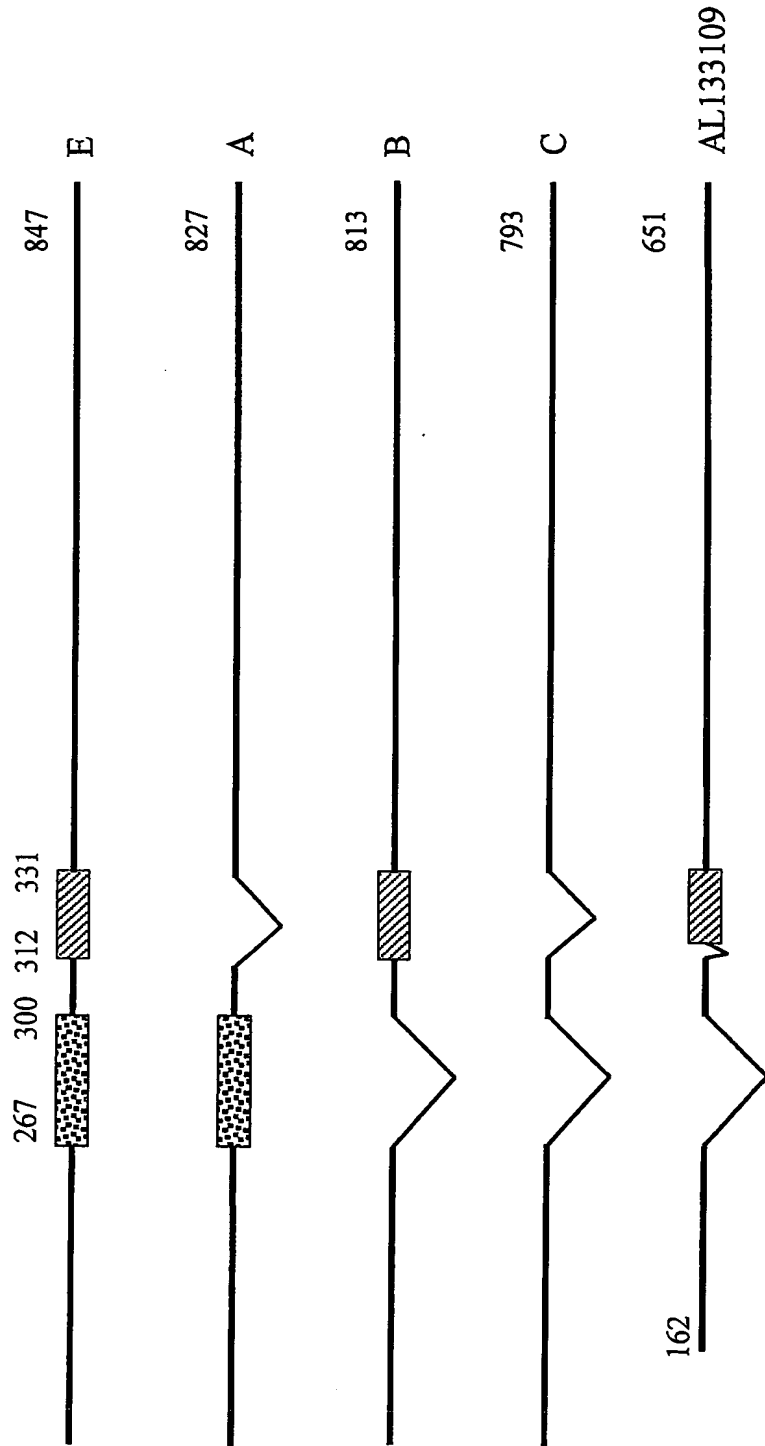
      720          *          740          *          760          *          780
BR137A : PVLSGQQGFQGLIGVQQPPQSQNVINNQQGTPVQSVMVSYPTMSSYQVPMTQGSQGLPQQSYQQP : 760
BR137C : PVLSGQQGFQGLIGVQQPPQSQNVINNQQGTPVQSVMVSYPTMSSYQVPMTQGSQGLPQQSYQQP : 726
BR137B : PVLSGQQGFQGLIGVQQPPQSQNVINNQQGTPVQSVMVSYPTMSSYQVPMTQGSQGLPQQSYQQP : 746
BR137E : PVLSGQQGFQGLIGVQQPPQSQNVINNQQGTPVQSVMVSYPTMSSYQVPMTQGSQGLPQQSYQQP : 780
AL133109 : PVLSGQQGFQGLIGVQQPPQSQNVINNQQGTPVQSVMVSYPTMSSYQVPMTQGSQGLPQQSYQQP : 584
BR137D : ----- : -

      *          800          *          820          *          840
BR137A : IMLPNQAGQGSLPATGMPVYCNVTPPTPQNNLRLIGPHCPSSSTVPVMSASCRTNCASMSNAGWQV : 825
BR137C : IMLPNQAGQGSLPATGMPVYCNVTPPTPQNNLRLIGPHCPSSSTVPVMSASCRTNCASMSNAGWQV : 791
BR137B : IMLPNQAGQGSLPATGMPVYCNVTPPTPQNNLRLIGPHCPSSSTVPVMSASCRTNCASMSNAGWQV : 811
BR137E : IMLPNQAGQGSLPATGMPVYCNVTPPTPQNNLRLIGPHCPSSSTVPVMSASCRTNCASMSNAGWQV : 845
AL133109 : IMLPNQAGQGSLPATGMPVYCNVTPPTPQNNLRLIGPHCPSSSTVPVMSASCRTNCASMSNAGWQV : 649
BR137D : ----- : -

BR137A : KF : 827
BR137C : KF : 793
BR137B : KF : 813
BR137E : KF : 847
AL133109 : KF : 651
BR137D : -- : -

```

FIG 10B



D/NM_016300

FIG 11

		*	20	*	40	*	60	
HOUSE :	POG	ITPT	LE	Q	EP			: 64
BR137A :								: 65
BR137C :								: 65
BR137B :								: 65
BR137E :								: 65
BR137D :								: 65

	*	80	*	100	*	120	*	
HOUSE :		SC	SH		PS		RE	: 129
BR137A :								: 130
BR137C :								: 130
BR137B :								: 130
BR137E :								: 130
BR137D :								: 89

	140	*	160	*	180	*	
HOUSE :							: 194
BR137A :							: 195
BR137C :							: 195
BR137B :							: 195
BR137E :							: 195
BR137D :							: -

	200	*	220	*	240	*	260	
HOUSE :								: 259
BR137A :								: 260
BR137C :								: 260
BR137B :								: 260
BR137E :								: 260
BR137D :								: -

	*	280	*	300	*	320	
HOUSE :		-NRMH	FRD	DRRS	KSI	EE	: 323
BR137A :		NRMH	FRD	DRRS	KSI	EE	: 311
BR137C :							: 277
BR137B :							: 291
BR137E :		NRMH	FRD	DRRS	KSI	EE	: 325
BR137D :							: -

FIG 12A

		*	340	*	360	*	380	*	
HOUSE	:	KKRQLF	AR	SS	TP	TP	TP	TP	: 388
BR137A	:	-----	-----	-----	-----	-----	-----	-----	: 370
BR137C	:	-----	-----	-----	-----	-----	-----	-----	: 336
BR137B	:	KKRQLF	AR	SS	TP	TP	TP	TP	: 356
BR137E	:	KKRQLF	AR	SS	TP	TP	TP	TP	: 390
BR137D	:	-----	-----	-----	-----	-----	-----	-----	: -
			400	*	420	*	440	*	
HOUSE	:	-----	-----	-----	-----	-----	-----	-----	: 451
BR137A	:	-----	-----	-----	-----	-----	-----	-----	: 435
BR137C	:	-----	-----	-----	-----	-----	-----	-----	: 401
BR137B	:	-----	-----	-----	-----	-----	-----	-----	: 421
BR137E	:	-----	-----	-----	-----	-----	-----	-----	: 455
BR137D	:	-----	-----	-----	-----	-----	-----	-----	: -
			460	*	480	*	500	*	520
MOUSE	:	-----	-----	-----	-----	-----	-----	-----	: 516
BR137A	:	-----	-----	-----	-----	-----	-----	-----	: 500
BR137C	:	-----	-----	-----	-----	-----	-----	-----	: 466
BR137B	:	-----	-----	-----	-----	-----	-----	-----	: 486
BR137E	:	-----	-----	-----	-----	-----	-----	-----	: 520
BR137D	:	-----	-----	-----	-----	-----	-----	-----	: -
			*	540	*	560	*	580	
MOUSE	:	-----	-----	-----	-----	-----	-----	-----	: 545
BR137A	:	-----	-----	-----	-----	-----	-----	-----	: 565
BR137C	:	-----	-----	-----	-----	-----	-----	-----	: 531
BR137B	:	-----	-----	-----	-----	-----	-----	-----	: 551
BR137E	:	-----	-----	-----	-----	-----	-----	-----	: 585
BR137D	:	-----	-----	-----	-----	-----	-----	-----	: -
		*	600	*	620	*	640	*	
MOUSE	:	EL	AM	SL	SM	TP	TP	TP	: 610
BR137A	:	-----	-----	-----	-----	-----	-----	-----	: 630
BR137C	:	-----	-----	-----	-----	-----	-----	-----	: 596
BR137B	:	-----	-----	-----	-----	-----	-----	-----	: 616
BR137E	:	-----	-----	-----	-----	-----	-----	-----	: 650
BR137D	:	-----	-----	-----	-----	-----	-----	-----	: -

FIG 12B

		*	340	*	360	*	380	*	
HOUSE	:	KKRQLF	ALPSS	ST	RPP	Q	T		: 388
BR137A	:	-----							: 370
BR137C	:	-----							: 336
BR137B	:	KKRQLF							: 356
BR137E	:	KKRQLF							: 390
BR137D	:	-----							: -
			400	*	420	*	440	*	
HOUSE	:		AS	T		HA	T	S	: 451
BR137A	:								: 435
BR137C	:								: 401
BR137B	:								: 421
BR137E	:								: 455
BR137D	:								: -
			460	*	480	*	500	*	520
HOUSE	:		P		S		G	T	: 516
BR137A	:								: 500
BR137C	:								: 466
BR137B	:								: 486
BR137E	:								: 520
BR137D	:								: -
			*	540	*	560	*	580	
HOUSE	:			AS					: 545
BR137A	:								: 565
BR137C	:								: 531
BR137B	:								: 551
BR137E	:								: 585
BR137D	:								: -
		*	600	*	620	*	640	*	
HOUSE	:	EL	AS	LSM		D			: 610
BR137A	:	LV		MTL					: 630
BR137C	:	LV		MTL					: 596
BR137B	:	LV		MTL					: 616
BR137E	:	LV		MTL					: 650
BR137D	:								: -

FIG 12C

16U 100 PCT.ST25
SEQUENCE LISTING

DT09 Rec'd PCT/PTO 01 OCT 2004

<110> OriGene Technologies, Inc

<120> NOVEL EXPRESSED GENES

<130> 16U 100 PCT

<150> US 10/112,372

<151> 2002-04-01

<150> US 60/382,614

<151> 2002-05-24

<150> US 10/164,717

<151> 2002-06-10

<150> US 10/167,631

<151> 2002-06-13

<150> US 10/177,917

<151> 2002-06-24

<150> US 60/399,125

<151> 2002-07-30

<160> 59

<170> PatentIn version 3.1

<210> 1

<211> 5536

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (242)..(646)

<223>

<400> 1

ggccccagtgt tgccagatct tatttttcaa cagaaactgg aaatgggggtt gttacatgaa 60

acctcccatc ctaccaacgt tggcagttta ttccagtta cttcctgaca tactaaggga 120

gctaacgaaa gcatgtctga aacaaagcat aactcggctc accggttttc cagtgttgac 180

ctgggggtact gaagcagata gtgtccatat atagatcctc accctctgca cttcggggcg 240

c atg gct gac ttc cag ctt cca gat agt att ctc tgg tgc caa aac cta 289

Met Ala Asp Phe Gln Leu Pro Asp Ser Ile Leu Trp Cys Gln Asn Leu

1

5

10

15

ttt tct ctg cct gtt tgg cag tct gga cat act aga gaa ttg atg ctc 337

Phe Ser Leu Pro Val Trp Gln Ser Gly His Thr Arg Glu Leu Met Leu

20

25

30

cag tgt tca gcc ttg agt gat ggg gaa ctg gtg tat aaa tat ccc agc 385

Gln Cys Ser Ala Leu Ser Asp Gly Glu Leu Val Tyr Lys Tyr Pro Ser

35

40

45

tcc ctc act cct tgg ttg agg tta act ctg ggg tgc atg ttc tac act 433

Ser Leu Thr Pro Trp Leu Arg Leu Thr Leu Gly Cys Met Phe Tyr Thr

50

55

60

ggg tcc cag ggt gtc ctc act gag att aag cat cca ctg ccc act gta 481

Gly Ser Gln Gly Val Leu Thr Glu Ile Lys His Pro Leu Pro Thr Val

65

70

75

80

ata gct ggt ttg ata atg cat ctt tta ttg tct ccc tct ctc ctc tgc 529

Ile Ala Gly Leu Ile Met His Leu Leu Leu Ser Pro Ser Leu Leu Cys

85

90

95

atc att tcc aca ctc cat tac aga ggt tcc ttg ccc tct caa att att 577

Ile Ile Ser Thr Leu His Tyr Arg Gly Ser Leu Pro Ser Gln Ile Ile

100

105

110

agc act cat ttt cca tct cga ctt cta aga atc cag att aga cag gca 625

Ser Thr His Phe Pro Ser Arg Leu Leu Arg Ile Gln Ile Arg Gln Ala

115	120	160 100 PCT.ST25 125	
tta ttt cat ttg gcc att aag tagatcttgt ggaagctgga ttttcatgcc			676
Leu Phe His Leu Ala Ile Lys			
130	135		
ataccccgaa agtaggcttt tatgtagaca tcatggaggg tgagggctga gatggaagaa			736
gaggtaaaat tggaccaagg aagagaaccc tgggtgaagg gttccagctc ttaaaagggg			796
gtcctgggta cctggagggc attattacca gatgacagag gatctggagt ggctcttgct			856
aataagtatc ttgggacaaa gagcagttgc atgcacagag agaaactccc aatgcatgaa			916
gaggagctct tcaaagatga attatgagag gcctattata taaataagga ggcaaaaaga			976
agcaaaggag aaccatcctg ttgtatcaat gtcggagggg ggtgactgtt tgcaccttat			1036
gtgccagaga gagtggctga ctaggaaggc aataccagag gatggaaggg cagaggcaag			1096
accatgggag gccccttttt cagcccatca gatgtcaca actctaattgt ctctctgca			1156
cttccacggt cagacctctc ccatcctgtc ttgggcttcc ctcaattgta tcttctttcc			1216
tcctcttctt gtggagtga tggagagctc tgcagaagggt ggagtctggg gtttaggaga			1276
ccattaaact atctgaatat ctctgatgat gactttgtga aaatgctcct accacctgga			1336
aggataaaca gagcacatca agatttgtaa agacaatacc aagtagagtt cagctgaaaa			1396
gaagcaggaa tgaatgtctt caggttacca cctcctctct gccagaggaa tcttctaagt			1456
aggccagatg gtagtaagacg atttactcac tccagatata cccgctaggg acatgatgtc			1516
atgctggttg tcccccttga gagggtgacc agcatggaga gcaactggca gatcaaaacc			1576
cccctgccct ccatttacta agctctgaat acagatgcag gactgcttca gccagaaaa			1636
aggtgtacta tctcttttct aatctttctg gccagaaagg gcaccttttt ctaattcttg			1696
acccataagg agcacctttt tttttcatth aaaaaattta attattatgg ctacataata			1756
gttgtatata ttacagggt atatgtgatg tttccataca ggcatacact atgtaaggat			1816
cagatcaggg taattgtgca tccatctcct caagcatgta tcatttcttt gtgttgga			1876
cattccaatt acactctttt agttatttta aaatataaga aaaattattg ctaattctcc			1936
cagagtgttg ggattacagg cgtgagccac catgctcggg cttagaagag aataatttga			1996
atgctcccag cataaagaaa ggataaatgt ttaagggtgat ggatatctat ctctattacc			2056
ctgatctgat cattacacat tatgtgaatg tatcgaaata tcacatatat tccaaaaata			2116
tgtacatcta ttatatatca ataaaaataa atatttgatt ctagagggtc tctcgagctc			2176
acacaccag ccgagctggt cttgagaccc caagcttcag ccagggtctt ccaagtttcc			2236
tagaacgatt ctcatgtgta tggagactgg tttcttctca aattcccaa gctggcagac			2296
accagctgg atccacctg ggcctgatca cccagttgtc tctacaataa atgaagtcag			2356
gcaggccac agaggcttcc aagggaccgg agagcaaact tgttcattag gcaaaattta			2416
ctgagtgcct actctgtgcc aggcattggt gaatgctggg gtgcaacatt gaaccagtca			2476
tggccctgcc ctcagcaagc tccaaacca gtcagagaga tggccaataa aataagcaac			2536
tactgtaccg aggaacaggg ctgtggcaaa caaatgacac tcatttttga atgagtgagt			2596
gaattcaaga ttcagtgaat gaatgaagtg cagtgtattg tgctatcctg gtctggagga			2656
tgagagaagg ctttctagaa gaggtgatag gacacctgga ggatgaataa atttagccat			2716
gtgaagtggg gaagagagga gtattctagg cagaggaaat ggcatgtgca aatggcctgg			2776
ggttagaaca agatgggatg gcagaaccaa gaaaactcga atctgaatca gtcagcttag			2836
gctgtggggg ccagggtggg atgggtggtg agagatgcta tttgtgaagt aggcaggacc			2896

16U 100 PCT.ST25

ggggtgtaaa aaacactgtc atccatgtca aagagttag atccattcaa agaaatggga	2956
tttttaaaca tgcaggagag gttggtattt tcaggcacca aatttaaccc aatgagaaca	3016
tttcaatagt gcctttatcc ctgttttctg gtgatgatgg aaaagcataa tgccttgtag	3076
atttctcagt tctgaccaca caagttacat gtggataagt cagagccagg tggtgatact	3136
ctgaaagtat ccctgtgagc tcagagtgtt gggttgagaa gatgaacaag gctagatcca	3196
cttctatata cactagccca gaggggcctc acattcaaga aatccttgcc tacggggatc	3256
catggtgcac caggaaaaat gaaattcgta gtcaagacta ggaacaagta cttgggagag	3316
aaaggcccag cctggcagcc tccaggtagc aaaccacaag ataatgaaa ttattgtact	3376
cacaggcccc tgcctttag tttctccagc ctactccct ccaggatggg cttgttgat	3436
tcacttgtaa tacactggct accaaccact caccatctct ggggcaaaga ttccatgatc	3496
ccatctttgt accaagccca agccagagat cgcagatctc aggggtctagt tgcctggcat	3556
gatacattcc tgacagaggt ggtggagata catgtgttac cattgaactc cagagttgag	3616
aaagatgtgg catctaaaat gcaaaaacag aagacaaata aataggagga ctgaatggga	3676
acagggcctt tacattattg ttttaaaagt ctactctag aaaactccac cagagcaacc	3736
ttattttgag gtcccttctc ccaaagcaaa gctatccagt gtgcatgca ggtaaagtgg	3796
cagctgtgct ggacagaact tgctaaagct agactatcac cttgcctgta tgcaaagcct	3856
cctgggagaa tctacaaagg aattcctgcc tcccagccc caggagctga catcctgacc	3916
gcaggagtta tgctattctc acagttccaa tgtgtctgtg tttcacctc atgaccaatt	3976
gctgagaagg cagaaatagt tcttaagaaa ttccagaagt agaagtcccc cctcctttag	4036
ggccaagatg cacaagataa agttggggga aagagtaggc tttttatggt gaaagctgcc	4096
ctaacatttg tggggtcaga gcaaaggatc agaagaagac ccacaaatta gatgtctaaa	4156
tgtttataaa taagtcagca aacttaatat gttctcttat cttggtcaac ataccttcaa	4216
aatgatctga aagttagtta tgaatttaga cccttgacc atttcagagt gccaggatat	4276
gacatggcag tggcgggaga atctggccct ggaacatcc ctttctctct ctgcatctgg	4336
cacatccaac accatgatgg ggcttacaca catgcatata gaccctcag cctgaacatc	4396
caagttctgt ccaaatagct acccttgac ctagaagtgt ggacacttga agcatggctc	4456
acctgcagga gaatggatcc atggatgagg tctacagaga cctggaaagt ggcataggag	4516
tcatttaaac agagaatttt gggatctcag gtacccagag agtggttaag gcgggcagca	4576
gtcttaaggg ggtgccttct tggccccca gatttctcat cccataggga tggttatctc	4636
tggagaaggc ccagggcagg ggtccctccc atccgcagct aagcatggta ctgctggcag	4696
aggccagcat caaacttaac tgtgtcatcc cagccacaac aggtctgaat atttaaagaa	4756
acccccgcc atcctgcttc taaacgcatg ccgccttcc caagtgtgg ctttccatcg	4816
cctcctttac tccctcattg gtcctagatc agccctagac agggttctcg gggccctct	4876
ttattcccag tcaccccaaa gcccaatttc caagcccctt ccaaagccct ctgttgcaag	4936
cgcacccctc ctcctcgtg cccctcaat attcacacct aggtagtaga tgcaatactc	4996
tagctgccac tggcatgttc cacaagggtt tcctgctcca acttggggcc taagggattg	5056
attaggttgg ctaggttagg tcccctttat gacaaaacca aaacagaatt gatgccccca	5116
cccgtaagg gcacttaaaa aacaaaact caaagttcag ggccaagttt gaggatgtgc	5176
agaaagggtg gtgttttttc ttgtcaattg cgactctaata gatggactca cgttgcccg	5236

16U 100 PCT.ST25

```

tcttcccttt ctcttacacc ttacctacct actaaaggag gagttcttgc ttggttaagtg 5296
gatataatcc gcaaagacat gagagaattt attagaagcc actcaagagc cttagctacc 5356
ttctacaagg ggaaaaggac acacacaaat atctatagtg accctttttt tcgtttattt 5416
ttggtcagac tgtttaactt ccattttttt tgtcccctcc tttcttttcc ccttttagttg 5476
aaaactgcta aaatgtcagt ttgcgacctt gactttatat ttaaaaaaaaa aaaaaaaaaa 5536

```

<210> 2
 <211> 135
 <212> PRT
 <213> Homo sapiens

<400> 2

```

Met Ala Asp Phe Gln Leu Pro Asp Ser Ile Leu Trp Cys Gln Asn Leu
1           5           10           15

```

```

Phe Ser Leu Pro Val Trp Gln Ser Gly His Thr Arg Glu Leu Met Leu
          20           25           30

```

```

Gln Cys Ser Ala Leu Ser Asp Gly Glu Leu Val Tyr Lys Tyr Pro Ser
          35           40           45

```

```

Ser Leu Thr Pro Trp Leu Arg Leu Thr Leu Gly Cys Met Phe Tyr Thr
          50           55           60

```

```

Gly Ser Gln Gly Val Leu Thr Glu Ile Lys His Pro Leu Pro Thr Val
65           70           75           80

```

```

Ile Ala Gly Leu Ile Met His Leu Leu Leu Ser Pro Ser Leu Leu Cys
          85           90           95

```

```

Ile Ile Ser Thr Leu His Tyr Arg Gly Ser Leu Pro Ser Gln Ile Ile
          100          105          110

```

```

Ser Thr His Phe Pro Ser Arg Leu Leu Arg Ile Gln Ile Arg Gln Ala
          115          120          125

```

```

Leu Phe His Leu Ala Ile Lys
130           135

```

<210> 3
 <211> 26
 <212> DNA
 <213> Homo sapiens

<400> 3

```

ggagctaacg aaagcatgtc tgaaac 26

```

<210> 4
 <211> 25
 <212> DNA
 <213> Homo sapiens

<400> 4

```

ggcaaggaac ctctgtaatg gagtg 25

```

<210> 5
 <211> 50
 <212> DNA
 <213> Homo sapiens

<400> 5

```

tgataatgta tctaactggg tcccgggtggg gatttctgag aacaggtggg 50

```

<210>	6	
<211>	50	
<212>	DNA	
<213>	Homo sapiens	
<400>	6	
aaaaacaata	aataacagtg	ccggcttccc
tgaggggctg	agagagactc	50
<210>	7	
<211>	50	
<212>	DNA	
<213>	Homo sapiens	
<400>	7	
aattaaattt	taaaatgtct	gagattgact
accctgaccc	agagagaaaag	50
<210>	8	
<211>	50	
<212>	DNA	
<213>	Homo sapiens	
<400>	8	
agaaagagtt	tataataact	ctgatgtctg
gatatggttt	aatggatccg	50
<210>	9	
<211>	50	
<212>	DNA	
<213>	Homo sapiens	
<400>	9	
ccaagaggga	ttaaaaaggt	cgagatggga
gagatggagc	aatacacttc	50
<210>	10	
<211>	50	
<212>	DNA	
<213>	Homo sapiens	
<400>	10	
gctcatgccc	catctaaaag	gggcagctgt
aattgctctc	agctgattct	50
<210>	11	
<211>	2270	
<212>	DNA	
<213>	Homo sapiens	
<220>		
<221>	CDS	
<222>	(157)..(1080)	
<223>		
<400>	11	
agacagcaga	gaggctgccc	tgctgcaatg
tcaccgtcgt	cactgcctct	gcaggctgca
ggcacctgcc	actactgcag	aggactgagg
ggccttggcc	cagcagggac	cccagggcct
tggggggactg	tgtgagctgg	aaacgtggct
ggccag	atg	ggc agc acc
Met	Gly	Ser Thr Met Glu
1		5
ccc cct ggg ggt gcg tac ctg cac ctg ggc gcc gtg aca tcc cct gtg	222	
Pro Pro Gly Gly Ala Tyr Leu His Leu Gly Ala Val Thr Ser Pro Val		
10	15	20
ggc aca gcc cgc gtg ctg cag ctg gcc ttt ggc tgc act acc ttc agc	270	
Gly Thr Ala Arg Val Leu Gln Leu Ala Phe Gly Cys Thr Thr Phe Ser		
25	30	35
ctg gtg gct cac cgg ggt ggc ttt gcg ggc gtc cag ggc acc ttc tgc	318	
Leu Val Ala His Arg Gly Gly Phe Ala Gly Val Gln Gly Thr Phe Cys		
40	45	50
atg gcc gcc tgg ggc ttc tgc ttc gcc gtc tct acg ctg gtg gtg gcc	360	

16U 100 PCT.ST25														
Met	Ala	Ala	Trp	Gly	Phe	Cys	Phe	Ala	Val	Ser	Ala	Leu	Val	Ala
55					60					65				70
tgt	gag	ttc	aca	cgg	ctc	cac	ggc	tgc	ctg	cgg	ctc	tcc	tgg	ggc
Cys	Glu	Phe	Thr	Arg	Leu	His	Gly	Cys	Leu	Arg	Leu	Ser	Trp	Gly
				75					80					85
ttc	acc	gcc	gcc	ttc	gcc	atg	ctg	gcc	acc	ctg	cta	tgc	gcg	acg
Phe	Thr	Ala	Ala	Phe	Ala	Met	Leu	Ala	Thr	Leu	Leu	Cys	Ala	Thr
			90					95					100	
gcg	gtc	ctg	tat	ccg	ctg	tac	ttt	gcc	cgg	cgg	gag	tgt	ccc	ccc
Ala	Val		Tyr	Pro	Leu	Tyr		Phe	Ala	Arg	Arg	Glu	Cys	Pro
		105					110					115		
ccc	gcc	ggc	tgt	gct	gcc	agg	gac	ttc	cgc	ctg	gca	gcc	agt	gtc
Pro	Ala	Gly	Cys	Ala	Ala	Arg	Asp	Phe	Arg	Leu	Ala	Ala	Ser	Val
		120				125					130			
gcc	ggg	ctc	ctc	ttc	ctg	gcc	tac	gct	gtg	gag	gtg	gcc	ctg	acg
Ala	Gly	Leu	Leu	Phe	Leu	Ala	Tyr	Ala	Val	Glu	Val	Ala	Leu	Thr
		135			140					145				150
gcc	cgg	ccc	ggc	cag	gtg	agc	agc	tat	atg	gcc	acg	gtg	tcg	ggg
Ala	Arg	Pro	Gly	Gln	Val	Ser	Ser	Tyr	Met	Ala	Thr	Val	Ser	Gly
				155					160					165
ctc	aag	atc	gtc	cag	gcc	ttc	gtg	gcc	tgc	atc	atc	ttc	ggg	gcg
Leu	Lys	Ile	Val	Gln	Ala	Phe	Val	Ala	Cys	Ile	Ile	Phe	Gly	Ala
			170					175					180	
gtc	cat	gac	agc	cgc	tac	ggg	cgc	tac	gtg	gcc	acc	cag	tgg	tgc
Val	His	Asp	Ser	Arg	Tyr	Gly	Arg	Tyr	Val	Ala	Thr	Gln	Trp	Cys
		185					190					195		
gcc	gtc	tac	agc	ctg	tgc	ttc	ctg	gcc	aca	gtg	gcc	gtg	gtg	gcc
Ala	Val	Tyr	Ser	Leu	Cys	Phe	Leu	Ala	Thr	Val	Ala	Val	Val	Ala
		200				205					210			
agt	gtg	atg	ggc	cac	aca	ggg	ggc	ctg	ggc	tgc	ccc	ttt	gac	cgg
Ser	Val	Met	Gly	His	Thr	Gly	Gly	Leu	Gly	Cys	Pro	Phe	Asp	Arg
		215			220					225				230
gtg	gtg	gtg	tac	acc	ttc	ctg	gct	gtg	ctc	ctg	tac	ctc	agc	gcc
Val	Val	Val	Tyr	Thr	Phe	Leu	Ala	Val	Leu	Leu	Tyr	Leu	Ser	Ala
				235					240					245
gtg	atc	tgg	cca	gtc	ttc	tgt	ttc	gat	ccc	aag	tac	ggt	gag	ccc
Val	Ile	Trp	Pro	Val	Phe	Cys	Phe	Asp	Pro	Lys	Tyr	Gly	Glu	Pro
			250					255					260	
cgg	ccc	ccc	aac	tgt	gct	cgg	ggc	agc	tgt	ccc	tgg	gac	agc	cag
Arg	Pro	Pro	Asn	Cys	Ala	Arg	Gly	Ser	Cys	Pro	Trp	Asp	Ser	Gln
			265				270					275		
gtg	gtg	gcc	atc	ttc	acc	tac	gtc	aac	ctg	ctc	ctg	tac	gtc	gtt
Val	Val	Ala	Ile	Phe	Thr	Tyr	Val	Asn	Leu	Leu	Leu	Tyr	Val	Val
		280				285					290			
ctc	gcc	tac	tcc	cag	agg	att	cgc	ttc	gtg	ccc	agc	ctg	tag	
Leu	Ala	Tyr	Ser	Gln	Arg	Ile	Arg	Phe	Val	Pro	Ser	Leu		
				300						305				
cccacagtgg	cagccccacc	acctctgctc	tctggccacc	agctccaggt	tctgcgaggg									
aaactcaggt	gaaccaaagt	gtcaagtagc	gagcaggagg	gaagccgcgg	gctcagggga									
gatgaagtac	aacgggagag	gtcgtcctccc	gacttccaga	ctgggctcag	agggacaagg									
gcaggacagt	agggcagagg	ccctgagcag	acgcaggggga	gagcctcacc	taggccaatc									
cccaactcac	tctctctcac	catttcccag	cgctcctggg	gctcagtgtc	ggtcactgga									
acatcataag	ggaaaacttt	aaagcagctg	ctgttgtgac	ccattttgca	gctggggacg									
tggaggctag	aagcatttgc	gaggcaccga	gtacctccat	tagtgtctgg	agaccacagt									
caggctgggt	aaggcaagct	gaaagaggaa	tgacttccag	aggagacgaa	cccaggagag									

16U 100 PCT.ST25

tagctgggccc aggcaggggg agggaggggc gtggggagcc ggaaccactc ctgggttcaga	1620
gaggcagctc ccagaaccct gggaagagag ccctgcctgg ggagctgtgc cagctcacag	1680
agaaggagcc agaaatctgg accagcctct cccaccacg tccgggaagc ctctgggctg	1740
gtagccacag tgtttggaag caggcttcct cgggagtctc cgcaggctgg aggtggcgcg	1800
ggtggcctgg aagagcctgc agagcccgcc ggcaccgagt gcacagtggg cggggaggac	1860
ctggttccgc catagccacg gagccccacc tggacactca ccattggctg tgacgaagca	1920
gcttcagcag tgcccggcgg gcatctgtgc actgtgggca tctgtggcac tgggaggag	1980
cccggctgag ggcggccgct ggacacagaa tctgggtact gctgcctct gctcaagggt	2040
ccagttgccg aaactcctga cgccggggcc atcctcctcc aggtccagc cagcttctcc	2100
tgcacagaag cccagcctgg tccagccagg agctgaccca ctggccaccc ctgagtccaa	2160
gccgggtggg cagtggcaca acagcccctc agcccattga ctgggccccca ttgacgtcct	2220
tgagcaggaa ataatgctg acatttatac gtaaaaaaaaa aaaaaaaaaa	2270

<210> 12
 <211> 307
 <212> PRT
 <213> Homo sapiens

<400> 12

Met Gly Ser Thr Met Glu Pro Pro Gly Gly Ala Tyr Leu His Leu Gly
 1 5 10 15

Ala Val Thr Ser Pro Val Gly Thr Ala Arg Val Leu Gln Leu Ala Phe
 20 25 30

Gly Cys Thr Thr Phe Ser Leu Val Ala His Arg Gly Gly Phe Ala Gly
 35 40 45

Val Gln Gly Thr Phe Cys Met Ala Ala Trp Gly Phe Cys Phe Ala Val
 50 55 60

Ser Ala Leu Val Val Ala Cys Glu Phe Thr Arg Leu His Gly Cys Leu
 65 70 75 80

Arg Leu Ser Trp Gly Asn Phe Thr Ala Ala Phe Ala Met Leu Ala Thr
 85 90 95

Leu Leu Cys Ala Thr Ala Ala Val Leu Tyr Pro Leu Tyr Phe Ala Arg
 100 105 110

Arg Glu Cys Pro Pro Glu Pro Ala Gly Cys Ala Ala Arg Asp Phe Arg
 115 120 125

Leu Ala Ala Ser Val Phe Ala Gly Leu Leu Phe Leu Ala Tyr Ala Val
 130 135 140

Glu Val Ala Leu Thr Arg Ala Arg Pro Gly Gln Val Ser Ser Tyr Met
 145 150 155 160

Ala Thr Val Ser Gly Leu Leu Lys Ile Val Gln Ala Phe Val Ala Cys
 165 170 175

Ile Ile Phe Gly Ala Leu Val His Asp Ser Arg Tyr Gly Arg Tyr Val
 180 185 190

16U 100 PCT.ST25

Ala Thr Gln Trp Cys Val Ala Val Tyr Ser Leu Cys Phe Leu Ala Thr
195 200 205

Val Ala Val Val Ala Leu Ser Val Met Gly His Thr Gly Gly Leu Gly
210 215 220

Cys Pro Phe Asp Arg Leu Val Val Val Tyr Thr Phe Leu Ala Val Leu
225 230 235 240

Leu Tyr Leu Ser Ala Ala Val Ile Trp Pro Val Phe Cys Phe Asp Pro
245 250 255

Lys Tyr Gly Glu Pro Lys Arg Pro Pro Asn Cys Ala Arg Gly Ser Cys
260 265 270

Pro Trp Asp Ser Gln Leu Val Val Ala Ile Phe Thr Tyr Val Asn Leu
275 280 285

Leu Leu Tyr Val Val Asp Leu Ala Tyr Ser Gln Arg Ile Arg Phe Val
290 295 300

Pro Ser Leu
305

<210> 13
<211> 24
<212> DNA
<213> Homo sapiens

<400> 13
actgcagagg actgaggggc cttg

24

<210> 14
<211> 24
<212> DNA
<213> Homo sapiens

<400> 14
aagacactgg ttgccaggcg gaag

24

<210> 15
<211> 153
<212> PRT
<213> Mus musculus

<400> 15

Met Gly Ser Thr Met Glu Pro Pro Gly Gly Ala Tyr Leu His Leu Gly
1 5 10 15

Ala Val Thr Ser Pro Val Gly Thr Ala Arg Met Leu Gln Leu Ala Phe
20 25 30

Gly Cys Thr Thr Phe Ser Leu Val Ala His Arg Gly Gly Phe Gly Gly
35 40 45

Val Gln Gly Thr Phe Cys Met Ala Ala Trp Gly Phe Cys Phe Ala Phe
50 55 60

Ser Val Leu Val Val Ala Cys Glu Phe Thr Lys Leu His Ser Cys Leu
65 70 75 80

Arg Leu Ser Trp Gly Asn Phe Thr Ala Ala Phe Ala Met Leu Ala Thr

85 90 160 100 PCT.ST25 95

Leu Leu Cys Ala Thr Ala Ala Val Ile Tyr Pro Leu Tyr Phe Thr Arg
100 105 110

Leu Glu Cys Pro Pro Glu Pro Ala Gly Cys Met Val Ala Pro Cys Gln
115 120 125

Arg Pro Ala Pro Glu Ser Pro Trp Lys Asp Asp Val Met Thr Ala
130 135 140

Met Glu Tyr Leu Ser Arg His Pro Thr
145 150

<210> 16
<211> 6455
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (431)..(5554)
<223>

<400> 16
gctttgtgca agaaagtgca agtttcccg tctggcttca tttttgttcc cttttgcaat 60
cctcctggct ccccccaaaa ccaagctagc aaagcaatgg ccccggttcc cccccaacgc 120
ctgacctgcg tttactggga ggagagcggg agaggggagcg cgcattctgg agcaggctgc 180
tttgactccg accacaggct gttttgtgca ggctgtccct cttcttcaaa atcgtgcatc 240
ccctccccga agcagcaggc agtgtgcctc cattcagcca catttggtat gcatgagcac 300
ggctgcagag agaggggagg tggctgtttt aagaaggttc aggggctcag gcaaggctac 360
ttgactagtc ttccaagttc caggaagcct ctgccotaat ggaatttgca ggtgtggaga 420
tgaccatggg atg cca gag ccg tgg ggg acc gtt tat ttt cta ggc att 469
Met Pro Glu Pro Trp Gly Thr Val Tyr Phe Leu Gly Ile
1 5 10

gct cag gtt ttc agt ttc ttg ttt tcc tgg tgg aat ttg gaa ggg gtc 517
Ala Gln Val Phe Ser Phe Leu Phe Ser Trp Trp Asn Leu Glu Gly Val
15 20 25

atg aat cag gct gat gct cct cga ccc cta aac tgg acc atc cgg aag 565
Met Asn Gln Ala Asp Ala Pro Arg Pro Leu Asn Trp Thr Ile Arg Lys
30 35 40 45

ctg tgc cac gca gcc ttt ctt cca tct gtc aga ctt ctg aag gct cag 613
Leu Cys His Ala Ala Phe Leu Pro Ser Val Arg Leu Leu Lys Ala Gln
50 55 60

aaa tcc tgg ata gaa aga gca ttt tat aaa aga gaa tgt gtc cac atc 661
Lys Ser Trp Ile Glu Arg Ala Phe Tyr Lys Arg Glu Cys Val His Ile
65 70 75

ata ccc agc acc aaa gac ccc cat agg tgt tgc tgt ggg cgt ctg ata 709
Ile Pro Ser Thr Lys Asp Pro His Arg Cys Cys Cys Gly Arg Leu Ile
80 85 90

ggc cag cat gtt ggc ctc acc ccc agt atc tcc gtg ctt cag aat gag 757
Gly Gln His Val Gly Leu Thr Pro Ser Ile Ser Val Leu Gln Asn Glu
95 100 105

aaa aat gaa agt cgc ctc tcc cga aat gac atc cag tct gaa aag tgg 805
Lys Asn Glu Ser Arg Leu Ser Arg Asn Asp Ile Gln Ser Glu Lys Trp
110 115 120 125

tcc atc agc aaa cac act caa ctc agc cct acg gat gct ttt ggg acc 853
Ser Ile Ser Lys His Thr Gln Leu Ser Pro Thr Asp Ala Phe Gly Thr
130 135 140

16U 100 PCT.ST25

att gag ttc caa gga ggt ggc cat tcc aac aaa gcc atg tat gtg cga Ile Glu Phe Gln Gly Gly Gly His Ser Asn Lys Ala Met Tyr Val Arg 145 150 155	901
gta tct ttt gat aca aaa cct gat ctc ctc tta cac ctg atg acc aag Val Ser Phe Asp Thr Lys Pro Asp Leu Leu Leu His Met Thr Lys 160 165 170	949
gaa tgg cag ttg gag ctt ccc aag ctt ctc atc tct gtc cat ggg ggc Glu Trp Gln Leu Glu Leu Pro Lys Leu Leu Ile Ser Val His Gly Gly 175 180 185	997
ctg cag aac ttt gaa ctc cag cca aaa ctc aag caa gtc ttt ggg aaa Leu Gln Asn Phe Glu Leu Gln Pro Lys Leu Lys Gln Val Phe Gly Lys 190 195 200 205	1045
ggg ctc atc aaa gca gca atg aca act gga gcg tgg ata ttc act gga Gly Leu Ile Lys Ala Ala Met Thr Thr Gly Ala Trp Ile Phe Thr Gly 210 215 220	1093
ggg gtt aac aca ggt gtt att cgt cat gtt ggc gat gcc ttg aag gat Gly Val Asn Thr Gly Val Ile Arg His Val Gly Asp Ala Leu Lys Asp 225 230 235	1141
cat gcc tct aag tct cga gga aag ata tgc acc ata ggt att gcc ccc His Ala Ser Lys Ser Arg Gly Lys Ile Cys Thr Ile Gly Ile Ala Pro 240 245 250	1189
tgg gga att gtg gaa aac cag gag gac ctc att gga aga gat gtt gtc Trp Gly Ile Val Glu Asn Gln Glu Asp Leu Ile Gly Arg Asp Val Val 255 260 265	1237
cgg cca tac cag acc atg tcc aat ccc atg agc aag ctc act gtt ctc Arg Pro Tyr Gln Thr Met Ser Asn Pro Met Ser Lys Leu Thr Val Leu 270 275 280 285	1285
aac agc atg cat tcc cac ttc att ctg gct gac aac ggg acc act gga Asn Ser Met His Ser His Phe Ile Leu Ala Asp Asn Gly Thr Thr Gly 290 295 300	1333
aaa tat gga gca gag gtg aaa ctt cga aga caa ctg gaa aag cat att Lys Tyr Gly Ala Glu Val Lys Leu Arg Arg Gln Leu Glu Lys His Ile 305 310 315	1381
tca ctc cag aag ata aac aca aga atc ggt caa ggt gtt cct gtg gtg Ser Leu Gln Lys Ile Asn Thr Arg Ile Gly Gln Gly Val Pro Val Val 320 325 330	1429
gca ctc ata gtg gaa gga gga ccc aat gtg atc tcg att gtt ttg gag Ala Leu Ile Val Glu Gly Gly Pro Asn Val Ile Ser Ile Val Leu Glu 335 340 345	1477
tac ctt cga gac acc cct ccc gtg cca gtg gtt gtc tgt gat ggg agt Tyr Leu Arg Asp Thr Pro Pro Val Pro Val Val Val Cys Asp Gly Ser 350 355 360 365	1525
gga cgg gca tcg gac atc ctg gcc ttt ggg cat aaa tac tca gaa gaa Gly Arg Ala Ser Asp Ile Leu Ala Phe Gly His Lys Tyr Ser Glu Glu 370 375 380	1573
ggc gga ctg ata aat gaa tct ttg agg gac cag ctg ttg gtg act ata Gly Gly Leu Ile Asn Glu Ser Leu Arg Asp Gln Leu Leu Val Thr Ile 385 390 395	1621
cag aag act ttc aca tac act cga acc caa gct cag cat ctg ttc atc Gln Lys Thr Phe Thr Tyr Thr Arg Thr Gln Ala Gln His Leu Phe Ile 400 405 410	1669
atc ctc atg gag tgc atg aag aag aag gaa ttg att acg gta ttt cgg Ile Leu Met Glu Cys Met Lys Lys Lys Glu Leu Ile Thr Val Phe Arg 415 420 425	1717
atg gga tca gaa gga cac cag gac att gat ttg gct atc ctg aca gct Met Gly Ser Glu Gly His Gln Asp Ile Asp Leu Ala Ile Leu Thr Ala 430 435 440 445	1765
tta ctc aaa gga gcc aat gcc tcg gcc cca gac caa ctg agc tta gct Leu Leu Lys Gly Ala Asn Ala Ser Ala Pro Asp Gln Leu Ser Leu Ala 1813	

															16U 100 PCT.ST25	
450					455					460						
tta	gcc	tgg	aac	aga	gtc	gac	atc	gct	cgc	agc	cag	atc	ttt	att	tac	1861
Leu	Ala	Trp	Asn	Arg	Val	Asp	Ile	Ala	Arg	Ser	Gln	Ile	Phe	Ile	Tyr	
465					470					475						
ggg	caa	cag	tgg	ccg	gtg	gga	tct	ctg	gag	caa	gcc	atg	ttg	gat	gcc	1909
Gly	Gln	Gln	Trp	Pro	Val	Gly	Ser	Leu	Glu	Gln	Ala	Met	Leu	Asp	Ala	
480					485					490						
tta	gtt	ctg	gac	aga	gtg	gat	ttt	gtg	aaa	tta	ctc	ata	gag	aat	gga	1957
Leu	Val	Leu	Asp	Arg	Val	Asp	Phe	Val	Lys	Leu	Leu	Ile	Glu	Asn	Gly	
495					500					505						
gta	agc	atg	cac	cgt	ttt	ctc	acc	atc	tcc	aga	cta	gag	gaa	ttg	tac	2005
Val	Ser	Met	His	Arg	Phe	Leu	Thr	Ile	Ser	Arg	Leu	Glu	Glu	Leu	Tyr	
510					515					520					525	
aat	acg	aga	cat	ggg	ccc	tca	aat	aca	ttg	tac	cac	ttg	gtc	agg	gat	2053
Asn	Thr	Arg	His	Gly	Pro	Ser	Asn	Thr	Leu	Tyr	His	Leu	Val	Arg	Asp	
530					535					540					545	
gtc	aaa	aag	ggg	aac	ctg	ccc	cca	gac	tac	aga	atc	agc	ctg	att	gac	2101
Val	Lys	Lys	Gly	Asn	Leu	Pro	Pro	Tyr	Arg	Ile	Ser	Leu	Ile	Asp		
545					550					555						
atc	ggc	ctg	gtg	atc	gag	tac	ctg	atg	ggc	ggg	gct	tat	cgc	tgc	aac	2149
Ile	Gly	Leu	Val	Ile	Glu	Tyr	Leu	Met	Gly	Gly	Ala	Tyr	Arg	Cys	Asn	
560					565					570						
tac	acg	cgc	aag	cgc	ttc	cgg	acc	ctc	tac	cac	aac	ctc	ttc	ggc	ccc	2197
Tyr	Thr	Arg	Lys	Arg	Phe	Arg	Thr	Leu	Tyr	His	Asn	Leu	Phe	Gly	Pro	
575					580					585						
aag	agg	ccc	aaa	gcc	ttg	aaa	ctg	ctg	gga	atg	gag	gat	gat	att	ccc	2245
Lys	Arg	Pro	Lys	Ala	Leu	Lys	Leu	Leu	Gly	Met	Glu	Asp	Asp	Ile	Pro	
590					595					600					605	
ttg	agg	cga	gga	aga	aag	aca	acc	aag	aaa	cgt	gaa	gaa	gag	gtg	gac	2293
Leu	Arg	Arg	Gly	Arg	Lys	Thr	Thr	Lys	Lys	Arg	Glu	Glu	Glu	Val	Asp	
610					615					620						
att	gac	ttg	gat	gat	cct	gag	atc	aac	cac	ttc	ccc	ttc	cct	ttc	cat	2341
Ile	Asp	Leu	Asp	Asp	Pro	Glu	Ile	Asn	His	Phe	Pro	Phe	Pro	Phe	His	
625					630					635						
gag	ctc	atg	gtg	tgg	gct	gtt	ctc	atg	aag	cgg	cag	aag	atg	gcc	ctg	2389
Glu	Leu	Met	Val	Trp	Ala	Val	Leu	Met	Lys	Arg	Gln	Lys	Met	Ala	Leu	
640					645					650						
ttc	ttc	tgg	cag	cac	ggt	gag	gag	gcc	atg	gcc	aag	gcc	ctg	gtg	gcc	2437
Phe	Phe	Trp	Gln	His	Gly	Glu	Glu	Ala	Met	Ala	Lys	Ala	Leu	Val	Ala	
655					660					665						
tgc	aag	ctc	tgc	aaa	gcc	atg	gct	cat	gag	gcc	tct	gag	aac	gac	atg	2485
Cys	Lys	Leu	Cys	Lys	Ala	Met	Ala	His	Glu	Ala	Ser	Glu	Asn	Asp	Met	
670					675					680					685	
gtt	gac	gac	att	tcc	cag	gag	ctg	aat	cac	aat	tcc	aga	gac	ttt	ggc	2533
Val	Asp	Asp	Ile	Ser	Gln	Glu	Leu	Asn	His	Asn	Ser	Arg	Asp	Phe	Gly	
690					695					700						
cag	ctg	gct	gtg	gag	ctc	ctg	gac	cag	tcc	tac	aag	cag	gac	gaa	cag	2581
Gln	Leu	Ala	Val	Glu	Leu	Leu	Asp	Gln	Ser	Tyr	Lys	Gln	Asp	Glu	Gln	
705					710					715						
ctg	gcc	atg	aaa	ctg	ctg	acg	tat	gag	ctg	aag	aac	tgg	agc	aac	gcc	2629
Leu	Ala	Met	Lys	Leu	Leu	Thr	Tyr	Glu	Leu	Lys	Asn	Trp	Ser	Asn	Ala	
720					725					730						
acg	tgc	ctg	cag	ctt	gcc	gtg	gct	gcc	aaa	cac	cgc	gac	ttc	atc	gcg	2677
Thr	Cys	Leu	Gln	Leu	Ala	Val	Ala	Ala	Lys	His	Arg	Asp	Phe	Ile	Ala	
735					740					745						
cac	acg	tgc	agc	cag	atg	ctg	ctc	acc	gac	atg	tgg	atg	ggc	cgg	ctc	2725
His	Thr	Cys	Ser	Gln	Met	Leu	Leu	Thr	Asp	Met	Trp	Met	Gly	Arg	Leu	
750					755					760					765	
cgc	atg	cgc	aag	aac	tca	ggc	ctc	aag	gta	att	ctg	gga	att	cta	ctt	2773

Page 12

											16U 100 PCT.ST25					
ccc Pro	tgc Cys	aag Lys	aca Thr	gga Gly 1085	gct Ala	tgg Trp	atc Ile	gtg Val	ccg Pro 1090	gcc Ala	atc Ile	atg Met	gcc Ala	tgc Cys 1095	3715	
tac Tyr	ctc Leu	tta Leu	gtg Val	gca Ala 1100	aac Asn	atc Ile	ttg Leu	ctg Leu	gtc Val 1105	aac Asn	ctc Leu	ctc Leu	att Ile	gct Ala 1110	3760	
gtc Val	ttt Phe	aac Asn	aat Asn	aca Thr 1115	ttt Phe	ttt Phe	gaa Glu	gta Val	aaa Lys 1120	tcg Ser	ata Ile	tcc Ser	aac Asn	caa Gln 1125	3805	
gtc Val	tgg Trp	aag Lys	ttt Phe	cag Gln 1130	agg Arg	tat Tyr	cag Gln	ctc Leu	atc Ile 1135	atg Met	act Thr	ttc Phe	cat His	gaa Glu 1140	3850	
agg Arg	cca Pro	gtt Val	ctg Leu	ccc Pro 1145	cca Pro	cca Pro	ctg Leu	atc Ile	atc Ile 1150	ttc Phe	agc Ser	cac His	atg Met	acc Thr 1155	3895	
atg Met	ata Ile	ttc Phe	cag Gln	cac His 1160	ctg Leu	tgc Cys	tgc Cys	cga Arg	tgg Trp 1165	agg Arg	aaa Lys	cac His	gag Glu	agc Ser 1170	3940	
gac Asp	ccg Pro	gat Asp	gaa Glu	agg Arg 1175	gac Asp	tac Tyr	ggc Gly	ctg Leu	aaa Lys 1180	ctc Leu	ttc Phe	ata Ile	acc Thr	gat Asp 1185	3985	
gat Asp	gag Glu	ctc Leu	aag Lys	aaa Lys 1190	gta Val	cat His	gac Asp	ttt Phe	gaa Glu 1195	gag Glu	caa Gln	tgc Cys	ata Ile	gaa Glu 1200	4030	
gaa Glu	tac Tyr	ttc Phe	aga Arg	gaa Glu 1205	aag Lys	gat Asp	gat Asp	cgg Arg	ttc Phe 1210	aac Asn	tca Ser	tct Ser	aat Asn	gat Asp 1215	4075	
gag Glu	agg Arg	ata Ile	cgg Arg	gtg Val 1220	act Thr	tca Ser	gaa Glu	agg Arg	gtg Val 1225	gag Glu	aac Asn	atg Met	tct Ser	atg Met 1230	4120	
cgg Arg	ctg Leu	gag Glu	gaa Glu	gtc Val 1235	aac Asn	gag Glu	aga Arg	gag Glu	cac His 1240	tcc Ser	atg Met	aag Lys	gct Ala	tca Ser 1245	4165	
ctc Leu	cag Gln	acc Thr	gtg Val	gac Asp 1250	atc Ile	cgg Arg	ctg Leu	gcg Ala	cag Gln 1255	ctg Leu	gaa Glu	gac Asp	ctt Leu	atc Ile 1260	4210	
ggg Gly	cgc Arg	atg Met	gcc Ala	acg Thr 1265	gcc Ala	ctg Leu	gag Glu	cgc Arg	ctg Leu 1270	aca Thr	ggg Gly	ctg Leu	gag Glu	cgg Arg 1275	4255	
gcc Ala	gag Glu	tcc Ser	aac Asn	aaa Lys 1280	atc Ile	cgc Arg	tcg Ser	agg Arg	acc Thr 1285	tcg Ser	tca Ser	gac Asp	tgc Cys	acg Thr 1290	4300	
gac Asp	gcc Ala	gcc Ala	tac Tyr	att Ile 1295	gtc Val	cgt Arg	cag Gln	agc Ser	agc Ser 1300	ttc Phe	aac Asn	agc Ser	cag Gln	gaa Glu 1305	4345	
ggg Gly	aac Asn	acc Thr	ttc Phe	aag Lys 1310	ctc Leu	caa Gln	gag Glu	agt Ser	ata Ile 1315	gac Asp	cct Pro	gca Ala	ggg Gly	gag Glu 1320	4390	
gag Glu	acc Thr	atg Met	tcc Ser	cca Pro 1325	act Thr	tct Ser	cca Pro	acc Thr	tta Leu 1330	atg Met	ccc Pro	cgt Arg	atg Met	cga Arg 1335	4435	
agc Ser	cat His	tct Ser	ttc Phe	tat Tyr 1340	tcg Ser	gtc Val	aat Asn	atg Met	aaa Lys 1345	gac Asp	aaa Lys	ggg Gly	ggg Gly	ata Ile 1350	4480	
gaa Glu	aag Lys	ttg Leu	gaa Glu	agt Ser 1355	att Ile	ttt Phe	aaa Lys	gaa Glu	agg Arg 1360	tcc Ser	ctg Leu	agc Ser	cta Leu	cac His 1365	4525	
cgg Arg	gct Ala	act Thr	agt Ser	tcc Ser 1370	cac His	tct Ser	gta Val	gca Ala	aaa Lys 1375	gaa Glu	ccc Pro	aaa Lys	gct Ala	cct Pro 1380	4570	

16U 100 PCT.ST25

gca gcc cct gcc aac acc ttg gcc att gtt cct gat tcc aga aga Ala Ala Pro Ala Asn Thr Leu Ala Ile Val Pro Asp Ser Arg Arg 1385 1390 1395	4615
cca tca tcg tgt ata gac atc tat gtc tct gct atg gat gag ctc Pro Ser Ser Cys Ile Asp Ile Tyr Val Ser Ala Met Asp Glu Leu 1400 1405 1410	4660
cac tgt gat ata gac cct ctg gac aat tcc gtg aac atc ctt ggg His Cys Asp Ile Asp Pro Leu Asp Asn Ser Val Asn Ile Leu Gly 1415 1420 1425	4705
ctg ggc gag cca agc ttt tca act cca gta cct tcc aca gcc cct Leu Gly Glu Pro Ser Phe Ser Thr Pro Val Pro Ser Thr Ala Pro 1430 1435 1440	4750
tca agt agt gcc tat gca aca ctt gca ccc aca gac aga cct cca Ser Ser Ser Ala Tyr Ala Thr Leu Ala Pro Thr Asp Arg Pro Pro 1445 1450 1455	4795
agc cgg agc att gat ttt gag gac atc acc tcc atg gac act aga Ser Arg Ser Ile Asp Phe Glu Asp Ile Thr Ser Met Asp Thr Arg 1460 1465 1470	4840
tct ttt tct tca gac tac acc cac ctg cca gaa tgc caa aac ccc Ser Phe Ser Ser Asp Tyr Thr His Leu Pro Glu Cys Gln Asn Pro 1475 1480 1485	4885
tgg gac tca gag cct ccg atg tac cac acc att gag cgt tcc aaa Trp Asp Ser Glu Pro Pro Met Tyr His Thr Ile Glu Arg Ser Lys 1490 1495 1500	4930
agt agc cgc tac cta gcc acc aca ccc ttt ctt cta gaa gag gct Ser Ser Arg Tyr Leu Ala Thr Thr Pro Phe Leu Leu Glu Glu Ala 1505 1510 1515	4975
ccc att gtg aaa tct cat agc ttt atg ttt tcc ccc tca agg agc Pro Ile Val Lys Ser His Ser Phe Met Phe Ser Pro Ser Arg Ser 1520 1525 1530	5020
tat tat gcc aac ttt ggg gtg cct gta aaa aca gca gaa tac aca Tyr Tyr Ala Asn Phe Gly Val Pro Val Lys Thr Ala Glu Tyr Thr 1535 1540 1545	5065
agt att aca gac tgt att gac aca agg tgt gtc aat gcc cct caa Ser Ile Thr Asp Cys Ile Asp Thr Arg Cys Val Asn Ala Pro Gln 1550 1555 1560	5110
gca att gcg gac aga gct gcc ttc cct gga ggt ctt gga gac aaa Ala Ile Ala Asp Arg Ala Ala Phe Pro Gly Gly Leu Gly Asp Lys 1565 1570 1575	5155
gtg gag gac tta act tgc tgc cat cca gag cga gaa gca gaa ctg Val Glu Asp Leu Thr Cys Cys His Pro Glu Arg Glu Ala Glu Leu 1580 1585 1590	5200
agt cac ccc agc tct gac agt gag gag aat gag gcc aaa ggc cgc Ser His Pro Ser Ser Asp Ser Glu Glu Asn Glu Ala Lys Gly Arg 1595 1600 1605	5245
aga gcc acc att gca ata tcc tcc cag gag ggt gat aac tca gag Arg Ala Thr Ile Ala Ile Ser Ser Gln Glu Gly Asp Asn Ser Glu 1610 1615 1620	5290
aga acc ctg tcc aac aac atc act gtt ccc aag ata gag cgc gcc Arg Thr Leu Ser Asn Asn Ile Thr Val Pro Lys Ile Glu Arg Ala 1625 1630 1635	5335
aac agc tac tcg gca gag gag cca agt gcg cca tat gca cac acc Asn Ser Tyr Ser Ala Glu Glu Pro Ser Ala Pro Tyr Ala His Thr 1640 1645 1650	5380
agg aag agc ttc tcc atc agt gac aaa ctg gac agg cag cgg aac Arg Lys Ser Phe Ser Ile Ser Asp Lys Leu Asp Arg Gln Arg Asn 1655 1660 1665	5425
aca gca agc ctg cga aat ccc ttc cag aga agc aag tcc tcc aag Thr Ala Ser Leu Arg Asn Pro Phe Gln Arg Ser Lys Ser Ser Lys 1670 1675 1680	5470

1670 1675 1680 16U 100 PCT.ST25

ccg gag ggc cga ggg gac agc ctg tcc atg agg aga ctg tcc aga 5515
 Pro Glu Gly Arg Gly Asp Ser Leu Ser Met Arg Arg Leu Ser Arg
 1685 1690 1695

aca tcg gct ttc caa agc ttt gaa agc aag cac aac taa accttcttaa 5564
 Thr Ser Ala Phe Gln Ser Phe Glu Ser Lys His Asn
 1700 1705

tatccgccac agaaggctca agaatccagc cctaaaattc tctccaactc cagtttttcc 5624
 cctttccttg aatcatacct gctttattct tagctgagca aaacaagcaa tgctttggga 5684
 ggtgttaact caaagggtgac ttctgggcca cagatcaaga aagcatttga tctgacccag 5744
 tgccagacac aggggattta aggcattgtc acacttgctg ggcagggagg gggaagagag 5804
 ggagaaggaa gggttagaga tgaatgtgta tccgcagtca cagcagaaag ccatgagagc 5864
 aggggaaaca aggggcttcg agcacgctcc atgccaggag gcactctgtg atttctgacc 5924
 attatcaaga gttgtaggat gcagggtctaa attgcaaaat aaaataaaat agccagcgta 5984
 cacaatgaga tattctaaac ttccattctg ttttcttttc acattggctc catcactggt 6044
 gactgatgaa gagcatcctc ttatttcagt ataagccggc agcaagcagt tctacctaac 6104
 gtccacatc cttctcatgc caacacttct gtaattgatc attataaaga aaaaacaagg 6164
 taacagtcac agttcacctg tctcttatct attcacttct ggtgccacaa ctgtttatcc 6224
 ttttttgaag aaaataaggg aacagaaatg cctttttgta ttgcaatcga aatgaaagaa 6284
 gagttgatgt taaaaaaca aaagtcaagt gatttattat atacagtggg cgttcaagtc 6344
 tagtcagca agctcaggag aatgtaatta aataatttta tattttttta tttattttgt 6404
 atctcacctg tcatggatga attcattcac tgaatatgta atattgaact t 6455

<210> 17
 <211> 1707
 <212> PRT
 <213> Homo sapiens

<400> 17

Met Pro Glu Pro Trp Gly Thr Val Tyr Phe Leu Gly Ile Ala Gln Val
 1 5 10 15

Phe Ser Phe Leu Phe Ser Trp Trp Asn Leu Glu Gly Val Met Asn Gln
 20 25 30

Ala Asp Ala Pro Arg Pro Leu Asn Trp Thr Ile Arg Lys Leu Cys His
 35 40 45

Ala Ala Phe Leu Pro Ser Val Arg Leu Leu Lys Ala Gln Lys Ser Trp
 50 55 60

Ile Glu Arg Ala Phe Tyr Lys Arg Glu Cys Val His Ile Ile Pro Ser
 65 70 75 80

Thr Lys Asp Pro His Arg Cys Cys Cys Gly Arg Leu Ile Gly Gln His
 85 90 95

Val Gly Leu Thr Pro Ser Ile Ser Val Leu Gln Asn Glu Lys Asn Glu
 100 105 110

Ser Arg Leu Ser Arg Asn Asp Ile Gln Ser Glu Lys Trp Ser Ile Ser
 115 120 125

16U 100 PCT.ST25

Lys His Thr Gln Leu Ser Pro Thr Asp Ala Phe Gly Thr Ile Glu Phe
 130 135 140
 Gln Gly Gly Gly His Ser Asn Lys Ala Met Tyr Val Arg Val Ser Phe
 145 150 155 160
 Asp Thr Lys Pro Asp Leu Leu Leu His Leu Met Thr Lys Glu Trp Gln
 165 170 175
 Leu Glu Leu Pro Lys Leu Leu Ile Ser Val His Gly Gly Leu Gln Asn
 180 185 190
 Phe Glu Leu Gln Pro Lys Leu Lys Gln Val Phe Gly Lys Gly Leu Ile
 195 200 205
 Lys Ala Ala Met Thr Thr Gly Ala Trp Ile Phe Thr Gly Gly Val Asn
 210 215 220
 Thr Gly Val Ile Arg His Val Gly Asp Ala Leu Lys Asp His Ala Ser
 225 230 235 240
 Lys Ser Arg Gly Lys Ile Cys Thr Ile Gly Ile Ala Pro Trp Gly Ile
 245 250 255
 Val Glu Asn Gln Glu Asp Leu Ile Gly Arg Asp Val Val Arg Pro Tyr
 260 265 270
 Gln Thr Met Ser Asn Pro Met Ser Lys Leu Thr Val Leu Asn Ser Met
 275 280 285
 His Ser His Phe Ile Leu Ala Asp Asn Gly Thr Thr Gly Lys Tyr Gly
 290 295 300
 Ala Glu Val Lys Leu Arg Arg Gln Leu Glu Lys His Ile Ser Leu Gln
 305 310 315 320
 Lys Ile Asn Thr Arg Ile Gly Gln Gly Val Pro Val Val Ala Leu Ile
 325 330 335
 Val Glu Gly Gly Pro Asn Val Ile Ser Ile Val Leu Glu Tyr Leu Arg
 340 345 350
 Asp Thr Pro Pro Val Pro Val Val Val Cys Asp Gly Ser Gly Arg Ala
 355 360 365
 Ser Asp Ile Leu Ala Phe Gly His Lys Tyr Ser Glu Glu Gly Gly Leu
 370 375 380
 Ile Asn Glu Ser Leu Arg Asp Gln Leu Leu Val Thr Ile Gln Lys Thr
 385 390 395 400
 Phe Thr Tyr Thr Arg Thr Gln Ala Gln His Leu Phe Ile Ile Leu Met
 405 410 415
 Glu Cys Met Lys Lys Lys Glu Leu Ile Thr Val Phe Arg Met Gly Ser
 420 425 430
 Glu Gly His Gln Asp Ile Asp Leu Ala Ile Leu Thr Ala Leu Leu Lys
 435 440 445

16U 100 PCT.ST25

Gly Ala Asn Ala Ser Ala Pro Asp Gln Leu Ser Leu Ala Leu Ala Trp
 450 455 460
 Asn Arg Val Asp Ile Ala Arg Ser Gln Ile Phe Ile Tyr Gly Gln Gln
 465 470 475 480
 Trp Pro Val Gly Ser Leu Glu Gln Ala Met Leu Asp Ala Leu Val Leu
 485 490 495
 Asp Arg Val Asp Phe Val Lys Leu Leu Ile Glu Asn Gly Val Ser Met
 500 505 510
 His Arg Phe Leu Thr Ile Ser Arg Leu Glu Glu Leu Tyr Asn Thr Arg
 515 520 525
 His Gly Pro Ser Asn Thr Leu Tyr His Leu Val Arg Asp Val Lys Lys
 530 535 540
 Gly Asn Leu Pro Pro Asp Tyr Arg Ile Ser Leu Ile Asp Ile Gly Leu
 545 550 555 560
 Val Ile Glu Tyr Leu Met Gly Gly Ala Tyr Arg Cys Asn Tyr Thr Arg
 565 570 575
 Lys Arg Phe Arg Thr Leu Tyr His Asn Leu Phe Gly Pro Lys Arg Pro
 580 585 590
 Lys Ala Leu Lys Leu Leu Gly Met Glu Asp Asp Ile Pro Leu Arg Arg
 595 600 605
 Gly Arg Lys Thr Thr Lys Lys Arg Glu Glu Glu Val Asp Ile Asp Leu
 610 615 620
 Asp Asp Pro Glu Ile Asn His Phe Pro Phe Pro Phe His Glu Leu Met
 625 630 635 640
 Val Trp Ala Val Leu Met Lys Arg Gln Lys Met Ala Leu Phe Phe Trp
 645 650 655
 Gln His Gly Glu Glu Ala Met Ala Lys Ala Leu Val Ala Cys Lys Leu
 660 665 670
 Cys Lys Ala Met Ala His Glu Ala Ser Glu Asn Asp Met Val Asp Asp
 675 680 685
 Ile Ser Gln Glu Leu Asn His Asn Ser Arg Asp Phe Gly Gln Leu Ala
 690 695 700
 Val Glu Leu Leu Asp Gln Ser Tyr Lys Gln Asp Glu Gln Leu Ala Met
 705 710 715 720
 Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ala Thr Cys Leu
 725 730 735
 Gln Leu Ala Val Ala Ala Lys His Arg Asp Phe Ile Ala His Thr Cys
 740 745 750
 Ser Gln Met Leu Leu Thr Asp Met Trp Met Gly Arg Leu Arg Met Arg

16U 100 PCT.ST25
 755 760 765
 Lys Asn Ser Gly Leu Lys Val Ile Leu Gly Ile Leu Leu Pro Pro Ser
 770 775 780
 Ile Leu Ser Leu Glu Phe Lys Asn Lys Asp Asp Met Pro Tyr Met Ser
 785 790 795 800
 Gln Ala Gln Glu Ile His Leu Gln Glu Lys Glu Ala Glu Glu Pro Glu
 805 810 815
 Lys Pro Thr Lys Glu Lys Glu Glu Glu Asp Met Glu Leu Thr Ala Met
 820 825 830
 Leu Gly Arg Asn Asn Gly Glu Ser Ser Arg Lys Lys Asp Glu Glu Glu
 835 840 845
 Val Gln Ser Lys His Arg Leu Ile Pro Leu Gly Arg Lys Ile Tyr Glu
 850 855 860
 Phe Tyr Asn Ala Pro Ile Val Lys Phe Trp Phe Tyr Thr Leu Ala Tyr
 865 870 875 880
 Ile Gly Tyr Leu Met Leu Phe Asn Tyr Ile Val Leu Val Lys Met Glu
 885 890 895
 Arg Trp Pro Ser Thr Gln Glu Trp Ile Val Ile Ser Tyr Ile Phe Thr
 900 905 910
 Leu Gly Ile Glu Lys Met Arg Glu Ile Leu Met Ser Glu Pro Gly Lys
 915 920 925
 Leu Leu Gln Lys Val Lys Val Trp Leu Gln Glu Tyr Trp Asn Val Thr
 930 935 940
 Asp Leu Ile Ala Ile Leu Leu Phe Ser Val Gly Met Ile Leu Arg Leu
 945 950 955 960
 Gln Asp Gln Pro Phe Arg Ser Asp Gly Arg Val Ile Tyr Cys Val Asn
 965 970 975
 Ile Ile Tyr Trp Tyr Ile Arg Leu Leu Asp Ile Phe Gly Val Asn Lys
 980 985 990
 Tyr Leu Gly Pro Tyr Val Met Met Ile Gly Lys Met Met Ile Asp Met
 995 1000 1005
 Met Tyr Phe Val Ile Ile Met Leu Val Val Leu Met Ser Phe Gly
 1010 1015 1020
 Val Ala Arg Gln Ala Ile Leu Phe Pro Asn Glu Glu Pro Ser Trp
 1025 1030 1035
 Lys Leu Ala Lys Asn Ile Phe Tyr Met Pro Tyr Trp Met Ile Tyr
 1040 1045 1050
 Gly Glu Val Phe Ala Asp Gln Ile Asp Pro Pro Cys Gly Gln Asn
 1055 1060 1065

Glu Thr Arg Glu Asp Gly Lys Ile Ile Gln Leu Pro Pro Cys Lys
 1070 1075 1080 16U 100 PCT.ST25

Thr Gly Ala Trp Ile Val Pro Ala Ile Met Ala Cys Tyr Leu Leu
 1085 1090 1095

Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Val Phe Asn
 1100 1105 1110

Asn Thr Phe Phe Glu Val Lys Ser Ile Ser Asn Gln Val Trp Lys
 1115 1120 1125

Phe Gln Arg Tyr Gln Leu Ile Met Thr Phe His Glu Arg Pro Val
 1130 1135 1140

Leu Pro Pro Pro Leu Ile Ile Phe Ser His Met Thr Met Ile Phe
 1145 1150 1155

Gln His Leu Cys Cys Arg Trp Arg Lys His Glu Ser Asp Pro Asp
 1160 1165 1170

Glu Arg Asp Tyr Gly Leu Lys Leu Phe Ile Thr Asp Asp Glu Leu
 1175 1180 1185

Lys Lys Val His Asp Phe Glu Glu Gln Cys Ile Glu Glu Tyr Phe
 1190 1195 1200

Arg Glu Lys Asp Asp Arg Phe Asn Ser Ser Asn Asp Glu Arg Ile
 1205 1210 1215

Arg Val Thr Ser Glu Arg Val Glu Asn Met Ser Met Arg Leu Glu
 1220 1225 1230

Glu Val Asn Glu Arg Glu His Ser Met Lys Ala Ser Leu Gln Thr
 1235 1240 1245

Val Asp Ile Arg Leu Ala Gln Leu Glu Asp Leu Ile Gly Arg Met
 1250 1255 1260

Ala Thr Ala Leu Glu Arg Leu Thr Gly Leu Glu Arg Ala Glu Ser
 1265 1270 1275

Asn Lys Ile Arg Ser Arg Thr Ser Ser Asp Cys Thr Asp Ala Ala
 1280 1285 1290

Tyr Ile Val Arg Gln Ser Ser Phe Asn Ser Gln Glu Gly Asn Thr
 1295 1300 1305

Phe Lys Leu Gln Glu Ser Ile Asp Pro Ala Gly Glu Glu Thr Met
 1310 1315 1320

Ser Pro Thr Ser Pro Thr Leu Met Pro Arg Met Arg Ser His Ser
 1325 1330 1335

Phe Tyr Ser Val Asn Met Lys Asp Lys Gly Gly Ile Glu Lys Leu
 1340 1345 1350

Glu Ser Ile Phe Lys Glu Arg Ser Leu Ser Leu His Arg Ala Thr
 1355 1360 1365

16U 100 PCT.ST25

Ser Ser His Ser Val Ala Lys Glu Pro Lys Ala Pro Ala Ala Pro
 1370 1375 1380
 Ala Asn Thr Leu Ala Ile Val Pro Asp Ser Arg Arg Pro Ser Ser
 1385 1390 1395
 Cys Ile Asp Ile Tyr Val Ser Ala Met Asp Glu Leu His Cys Asp
 1400 1405 1410
 Ile Asp Pro Leu Asp Asn Ser Val Asn Ile Leu Gly Leu Gly Glu
 1415 1420 1425
 Pro Ser Phe Ser Thr Pro Val Pro Ser Thr Ala Pro Ser Ser Ser
 1430 1435 1440
 Ala Tyr Ala Thr Leu Ala Pro Thr Asp Arg Pro Pro Ser Arg Ser
 1445 1450 1455
 Ile Asp Phe Glu Asp Ile Thr Ser Met Asp Thr Arg Ser Phe Ser
 1460 1465 1470
 Ser Asp Tyr Thr His Leu Pro Glu Cys Gln Asn Pro Trp Asp Ser
 1475 1480 1485
 Glu Pro Pro Met Tyr His Thr Ile Glu Arg Ser Lys Ser Ser Arg
 1490 1495 1500
 Tyr Leu Ala Thr Thr Pro Phe Leu Leu Glu Glu Ala Pro Ile Val
 1505 1510 1515
 Lys Ser His Ser Phe Met Phe Ser Pro Ser Arg Ser Tyr Tyr Ala
 1520 1525 1530
 Asn Phe Gly Val Pro Val Lys Thr Ala Glu Tyr Thr Ser Ile Thr
 1535 1540 1545
 Asp Cys Ile Asp Thr Arg Cys Val Asn Ala Pro Gln Ala Ile Ala
 1550 1555 1560
 Asp Arg Ala Ala Phe Pro Gly Gly Leu Gly Asp Lys Val Glu Asp
 1565 1570 1575
 Leu Thr Cys Cys His Pro Glu Arg Glu Ala Glu Leu Ser His Pro
 1580 1585 1590
 Ser Ser Asp Ser Glu Glu Asn Glu Ala Lys Gly Arg Arg Ala Thr
 1595 1600 1605
 Ile Ala Ile Ser Ser Gln Glu Gly Asp Asn Ser Glu Arg Thr Leu
 1610 1615 1620
 Ser Asn Asn Ile Thr Val Pro Lys Ile Glu Arg Ala Asn Ser Tyr
 1625 1630 1635
 Ser Ala Glu Glu Pro Ser Ala Pro Tyr Ala His Thr Arg Lys Ser
 1640 1645 1650
 Phe Ser Ile Ser Asp Lys Leu Asp Arg Gln Arg Asn Thr Ala Ser
 1655 1660 1665

16U 100 PCT.ST25

Leu Arg Asn Pro Phe Gln Arg Ser Lys Ser Ser Lys Pro Glu Gly
 1670 1675 1680

Arg Gly Asp Ser Leu Ser Met Arg Arg Leu Ser Arg Thr Ser Ala
 1685 1690 1695

Phe Gln Ser Phe Glu Ser Lys His Asn
 1700 1705

<210> 18
 <211> 988
 <212> PRT
 <213> Homo sapiens

<400> 18

Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ala Thr Cys
 1 5 10 15

Leu Gln Leu Ala Val Ala Ala Lys His Arg Asp Phe Ile Ala His Thr
 20 25 30

Cys Ser Gln Met Leu Leu Thr Asp Met Trp Met Gly Arg Leu Arg Met
 35 40 45

Arg Lys Asn Ser Gly Leu Lys Val Ile Leu Gly Ile Leu Leu Pro Pro
 50 55 60

Ser Ile Leu Ser Leu Glu Phe Lys Asn Lys Asp Asp Met Pro Tyr Met
 65 70 75 80

Ser Gln Ala Gln Glu Ile His Leu Gln Glu Lys Glu Ala Glu Glu Pro
 85 90 95

Glu Lys Pro Thr Lys Glu Lys Glu Glu Glu Asp Met Glu Leu Thr Ala
 100 105 110

Met Leu Gly Arg Asn Asn Gly Glu Ser Ser Arg Lys Lys Asp Glu Glu
 115 120 125

Glu Val Gln Ser Lys His Arg Leu Ile Pro Leu Gly Arg Lys Ile Tyr
 130 135 140

Glu Phe Tyr Asn Ala Pro Ile Val Lys Phe Trp Phe Tyr Thr Leu Ala
 145 150 155 160

Tyr Ile Gly Tyr Leu Met Leu Phe Asn Tyr Ile Val Leu Val Lys Met
 165 170 175

Glu Arg Trp Pro Ser Thr Gln Glu Trp Ile Val Ile Ser Tyr Ile Phe
 180 185 190

Thr Leu Gly Ile Glu Lys Met Arg Glu Ile Leu Met Ser Glu Pro Gly
 195 200 205

Lys Leu Leu Gln Lys Val Lys Val Trp Leu Gln Glu Tyr Trp Asn Val
 210 215 220

Thr Asp Leu Ile Ala Ile Leu Leu Phe Ser Val Gly Met Ile Leu Arg
 225 230 235 240

16U 100 PCT.ST25

Leu Gln Asp Gln Pro Phe Arg Ser Asp Gly Arg Val Ile Tyr Cys Val
 245 250 255
 Asn Ile Ile Tyr Trp Tyr Ile Arg Leu Leu Asp Ile Phe Gly Val Asn
 260 265 270
 Lys Tyr Leu Gly Pro Tyr Val Met Met Ile Gly Lys Met Met Ile Asp
 275 280 285
 Met Met Tyr Phe Val Ile Ile Met Leu Val Val Leu Met Ser Phe Gly
 290 295 300
 Val Ala Arg Gln Ala Ile Leu Phe Pro Asn Glu Glu Pro Ser Trp Lys
 305 310 315 320
 Leu Ala Lys Asn Ile Phe Tyr Met Pro Tyr Trp Met Ile Tyr Gly Glu
 325 330 335
 Val Phe Ala Asp Gln Ile Asp Pro Pro Cys Gly Gln Asn Glu Thr Arg
 340 345 350
 Glu Asp Gly Lys Ile Ile Gln Leu Pro Pro Cys Lys Thr Gly Ala Trp
 355 360 365
 Ile Val Pro Ala Ile Met Ala Cys Tyr Leu Leu Val Ala Asn Ile Leu
 370 375 380
 Leu Val Asn Leu Leu Ile Ala Val Phe Asn Asn Thr Phe Phe Glu Val
 385 390 395 400
 Lys Ser Ile Ser Asn Gln Val Trp Lys Phe Gln Arg Tyr Gln Leu Ile
 405 410 415
 Met Thr Phe His Glu Arg Pro Val Leu Pro Pro Pro Leu Ile Ile Phe
 420 425 430
 Ser His Met Thr Met Ile Phe Gln His Leu Cys Cys Arg Trp Arg Lys
 435 440 445
 His Glu Ser Asp Pro Asp Glu Arg Asp Tyr Gly Leu Lys Leu Phe Ile
 450 455 460
 Thr Asp Asp Glu Leu Lys Lys Val His Asp Phe Glu Glu Gln Cys Ile
 465 470 475 480
 Glu Glu Tyr Phe Arg Glu Lys Asp Asp Arg Phe Asn Ser Ser Asn Asp
 485 490 495
 Glu Arg Ile Arg Val Thr Ser Glu Arg Val Glu Asn Met Ser Met Arg
 500 505 510
 Leu Glu Glu Val Asn Glu Arg Glu His Ser Met Lys Ala Ser Leu Gln
 515 520 525
 Thr Val Asp Ile Arg Leu Ala Gln Leu Glu Asp Leu Ile Gly Arg Met
 530 535 540
 Ala Thr Ala Leu Glu Arg Leu Thr Gly Leu Glu Arg Ala Glu Ser Asn

Page 23

16U 100 PCT.ST25

Pro Glu Arg Glu Ala Glu Leu Ser His Pro Ser Ser Asp Ser Glu Glu
865 870 875 880

Asn Glu Ala Lys Gly Arg Arg Ala Thr Ile Ala Ile Ser Ser Gln Glu
885 890 895

Gly Asp Asn Ser Glu Arg Thr Leu Ser Asn Asn Ile Thr Val Pro Lys
900 905 910

Ile Glu Arg Ala Asn Ser Tyr Ser Ala Glu Glu Pro Ser Ala Pro Tyr
915 920 925

Ala His Thr Arg Lys Ser Phe Ser Ile Ser Asp Lys Leu Asp Arg Gln
930 935 940

Arg Asn Thr Ala Ser Leu Arg Asn Pro Phe Gln Arg Ser Lys Ser Ser
945 950 955 960

Lys Pro Glu Gly Arg Gly Asp Ser Leu Ser Met Arg Arg Leu Ser Arg
965 970 975

Thr Ser Ala Phe Gln Ser Phe Glu Ser Lys His Asn
980 985

<210> 19
<211> 1017
<212> PRT
<213> Homo sapiens

<400> 19

Gln Glu Leu Asn His Asn Ser Arg Asp Phe Gly Gln Leu Ala Val Glu
1 5 10 15

Leu Leu Asp Gln Ser Tyr Lys Gln Asp Glu Gln Leu Ala Met Lys Leu
20 25 30

Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ala Thr Cys Leu Gln Leu
35 40 45

Ala Val Ala Ala Lys His Arg Asp Phe Ile Ala His Thr Cys Ser Gln
50 55 60

Met Leu Leu Thr Asp Met Trp Met Gly Arg Leu Arg Met Arg Lys Asn
65 70 75 80

Ser Gly Leu Lys Val Ile Leu Gly Ile Leu Leu Pro Pro Ser Ile Leu
85 90 95

Ser Leu Glu Phe Lys Asn Lys Asp Asp Met Pro Tyr Met Ser Gln Ala
100 105 110

Gln Glu Ile His Leu Gln Glu Lys Glu Ala Glu Glu Pro Glu Lys Pro
115 120 125

Thr Lys Glu Lys Glu Glu Glu Asp Met Glu Leu Thr Ala Met Leu Gly
130 135 140

Arg Asn Asn Gly Glu Ser Ser Arg Lys Lys Asp Glu Glu Glu Val Gln
145 150 155 160

16U 100 PCT.ST25

Ser Lys His Arg Leu Ile Pro Leu Gly Arg Lys Ile Tyr Glu Phe Tyr
165 170 175

Asn Ala Pro Ile Val Lys Phe Trp Phe Tyr Thr Leu Ala Tyr Ile Gly
180 185 190

Tyr Leu Met Leu Phe Asn Tyr Ile Val Leu Val Lys Met Glu Arg Trp
195 200 205

Pro Ser Thr Gln Glu Trp Ile Val Ile Ser Tyr Ile Phe Thr Leu Gly
210 215 220

Ile Glu Lys Met Arg Glu Ile Leu Met Ser Glu Pro Gly Lys Leu Leu
225 230 235 240

Gln Lys Val Lys Val Trp Leu Gln Glu Tyr Trp Asn Val Thr Asp Leu
245 250 255

Ile Ala Ile Leu Leu Phe Ser Val Gly Met Ile Leu Arg Leu Gln Asp
260 265 270

Gln Pro Phe Arg Ser Asp Gly Arg Val Ile Tyr Cys Val Asn Ile Ile
275 280 285

Tyr Trp Tyr Ile Arg Leu Leu Asp Ile Phe Gly Val Asn Lys Tyr Leu
290 295 300

Gly Pro Tyr Val Met Met Ile Gly Lys Met Met Ile Asp Met Met Tyr
305 310 315 320

Phe Val Ile Ile Met Leu Val Val Leu Met Ser Phe Gly Val Ala Arg
325 330 335

Gln Ala Ile Leu Phe Pro Asn Glu Glu Pro Ser Trp Lys Leu Ala Lys
340 345 350

Asn Ile Phe Tyr Met Pro Tyr Trp Met Ile Tyr Gly Glu Val Phe Ala
355 360 365

Asp Gln Ile Asp Pro Pro Cys Gly Gln Asn Glu Thr Arg Glu Asp Gly
370 375 380

Lys Ile Ile Gln Leu Pro Pro Cys Lys Thr Gly Ala Trp Ile Val Pro
385 390 395 400

Ala Ile Met Ala Cys Tyr Leu Leu Val Ala Asn Ile Leu Leu Val Asn
405 410 415

Leu Leu Ile Ala Val Phe Asn Asn Thr Phe Phe Glu Val Lys Ser Ile
420 425 430

Ser Asn Gln Val Trp Lys Phe Gln Arg Tyr Gln Leu Ile Met Thr Phe
435 440 445

His Glu Arg Pro Val Leu Pro Pro Pro Leu Ile Ile Phe Ser His Met
450 455 460

Thr Met Ile Phe Gln His Leu Cys Cys Arg Trp Arg Lys His Glu Ser
465 470 475 480

16U 100 PCT.ST25

Asp Pro Asp Glu Arg Asp Tyr Gly Leu Lys Leu Phe Ile Thr Asp Asp
 485 490 495
 Glu Leu Lys Lys Val His Asp Phe Glu Glu Gln Cys Ile Glu Glu Tyr
 500 505 510
 Phe Arg Glu Lys Asp Asp Arg Phe Asn Ser Ser Asn Asp Glu Arg Ile
 515 520 525
 Arg Val Thr Ser Glu Arg Val Glu Asn Met Ser Met Arg Leu Glu Glu
 530 535 540
 Val Asn Glu Arg Glu His Ser Met Lys Ala Ser Leu Gln Thr Val Asp
 545 550 555 560
 Ile Arg Leu Ala Gln Leu Glu Asp Leu Ile Gly Arg Met Ala Thr Ala
 565 570 575
 Leu Glu Arg Leu Thr Gly Leu Glu Arg Ala Glu Ser Asn Lys Ile Arg
 580 585 590
 Ser Arg Thr Ser Ser Asp Cys Thr Asp Ala Ala Tyr Ile Val Arg Gln
 595 600 605
 Ser Ser Phe Asn Ser Gln Glu Gly Asn Thr Phe Lys Leu Gln Glu Ser
 610 615 620
 Ile Asp Pro Ala Gly Glu Glu Thr Met Ser Pro Thr Ser Pro Thr Leu
 625 630 635 640
 Met Pro Arg Met Arg Ser His Ser Phe Tyr Ser Val Asn Met Lys Asp
 645 650 655
 Lys Gly Gly Ile Glu Lys Leu Glu Ser Ile Phe Lys Glu Arg Ser Leu
 660 665 670
 Ser Leu His Arg Ala Thr Ser Ser His Ser Val Ala Lys Glu Pro Lys
 675 680 685
 Ala Pro Ala Ala Pro Ala Asn Thr Leu Ala Ile Val Pro Asp Ser Arg
 690 695 700
 Arg Pro Ser Ser Cys Ile Asp Ile Tyr Val Ser Ala Met Asp Glu Leu
 705 710 715 720
 His Cys Asp Ile Asp Pro Leu Asp Asn Ser Val Asn Ile Leu Gly Leu
 725 730 735
 Gly Glu Pro Ser Phe Ser Thr Pro Val Pro Ser Thr Ala Pro Ser Ser
 740 745 750
 Ser Ala Tyr Ala Thr Leu Ala Pro Thr Asp Arg Pro Pro Ser Arg Ser
 755 760 765
 Ile Asp Phe Glu Asp Ile Thr Ser Met Asp Thr Arg Ser Phe Ser Ser
 770 775 780
 Asp Tyr Thr His Leu Pro Glu Cys Gln Asn Pro Trp Asp Ser Glu Pro
 785 790 795 800

16U 100 PCT.ST25

Pro Met Tyr His Thr Ile Glu Arg Ser Lys Ser Ser Arg Tyr Leu Ala
 805 810 815
 Thr Thr Pro Phe Leu Leu Glu Glu Ala Pro Ile Val Lys Ser His Ser
 820 825 830
 Phe Met Phe Ser Pro Ser Arg Ser Tyr Tyr Ala Asn Phe Gly Val Pro
 835 840 845
 Val Lys Thr Ala Glu Tyr Thr Ser Ile Thr Asp Cys Ile Asp Thr Arg
 850 855 860
 Cys Val Asn Ala Pro Gln Ala Ile Ala Asp Arg Ala Ala Phe Pro Gly
 865 870 875 880
 Gly Leu Gly Asp Lys Val Glu Asp Leu Thr Cys Cys His Pro Glu Arg
 885 890 895
 Glu Ala Glu Leu Ser His Pro Ser Ser Asp Ser Glu Glu Asn Glu Ala
 900 905 910
 Lys Gly Arg Arg Ala Thr Ile Ala Ile Ser Ser Gln Glu Gly Asp Asn
 915 920 925
 Ser Glu Arg Thr Leu Ser Asn Asn Ile Thr Val Pro Lys Ile Glu Arg
 930 935 940
 Ala Asn Ser Tyr Ser Ala Glu Glu Pro Ser Ala Pro Tyr Ala His Thr
 945 950 955 960
 Arg Lys Ser Phe Ser Ile Ser Asp Lys Leu Asp Arg Gln Arg Asn Thr
 965 970 975
 Ala Ser Leu Arg Asn Pro Phe Gln Arg Ser Lys Ser Ser Lys Pro Glu
 980 985 990
 Gly Arg Gly Asp Ser Leu Ser Met Arg Lys Leu Ser Arg Thr Ser Ala
 995 1000 1005
 Phe Gln Ser Phe Glu Ser Lys His Thr
 1010 1015
 <210> 20
 <211> 736
 <212> PRT
 <213> Mus musculus
 <400> 20
 Met Val Leu Gly Thr Gly Thr Phe Leu Ser Ser Gln His Thr Ala Gly
 1 5 10 15
 Arg Leu Pro Pro Gly Ala Phe Ala Lys Gln Arg Leu Leu Cys Gly Ala
 20 25 30
 Ala Leu Leu Leu Tyr Val Ser Ala Asn Asn Pro Ile Gln Ala Gln Ser
 35 40 45
 Val Pro Ile Met Leu Ser Gln Arg Gly Leu Leu Ala Thr Cys Thr His
 50 55 60

16U 100 PCT.ST25

Ser Gly Val Phe Leu Leu Pro Tyr Arg Leu Pro Pro Tyr Thr Gln Leu
 65 70 75 80
 Ala Pro Cys Gly Gln Asn Glu Thr Arg Glu Asp Gly Lys Thr Ile Gln
 85 90 95
 Leu Pro Pro Cys Lys Thr Gly Ala Trp Ile Val Pro Ala Ile Met Ala
 100 105 110
 Cys Tyr Leu Leu Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala
 115 120 125
 Val Phe Asn Asn Thr Phe Phe Glu Val Lys Ser Ile Ser Asn Gln Val
 130 135 140
 Trp Lys Phe Gln Arg Tyr Gln Leu Ile Met Thr Phe His Glu Arg Pro
 145 150 155 160
 Val Leu Pro Pro Pro Leu Ile Ile Phe Ser His Met Thr Met Ile Phe
 165 170 175
 Gln His Val Cys Cys Arg Trp Arg Lys His Glu Ser Asp Gln Asp Glu
 180 185 190
 Arg Asp Tyr Gly Leu Lys Phe Leu Ile Thr Gly Asp Glu Leu Arg Lys
 195 200 205
 Val His Asp Phe Glu Glu Gln Cys Ile Glu Glu Tyr Phe Arg Glu Lys
 210 215 220
 Asp Asp Arg Phe Asn Ser Ser Asn Asp Glu Arg Ile Arg Val Thr Ser
 225 230 235 240
 Glu Arg Val Glu Asn Met Ser Met Arg Leu Glu Glu Val Asn Glu Arg
 245 250 255
 Glu His Ser Met Lys Ala Ser Leu Gln Thr Val Asp Ile Arg Leu Ala
 260 265 270
 Gln Leu Glu Asp Leu Ile Gly Arg Met Ala Thr Ala Leu Glu Arg Leu
 275 280 285
 Thr Gly Leu Glu Arg Ala Glu Ser Asn Lys Ile Arg Ser Arg Thr Ser
 290 295 300
 Ser Asp Cys Thr Asp Ala Ala Tyr Ile Val Arg Gln Ser Ser Phe Asn
 305 310 315 320
 Ser Gln Glu Gly Asn Thr Phe Lys Leu Gln Glu Ser Ile Asp Pro Ala
 325 330 335
 Gly Glu Glu Thr Ile Ser Pro Thr Ser Pro Thr Leu Met Pro Arg Met
 340 345 350
 Arg Ser His Ser Phe Tyr Ser Val Asn Val Lys Asp Lys Gly Gly Ile
 355 360 365
 Glu Lys Leu Glu Ser Ile Phe Lys Glu Arg Ser Leu Ser Leu His Arg

370 375 160 100 PCT.ST25
 380
 Ala Thr Ser Ser His Ser Val Ala Lys Glu Pro Lys Ala Pro Ala Ala
 385 390 395 400
 Pro Ala Asn Thr Leu Ala Ile Val Pro Asp Ser Arg Arg Pro Ser Ser
 405 410 415
 Cys Ile Asp Ile Tyr Val Ser Ala Met Asp Glu Leu His Cys Asp Ile
 420 425 430
 Glu Pro Leu Asp Asn Ser Met Asn Ile Leu Gly Leu Gly Glu Pro Ser
 435 440 445
 Phe Ser Ala Leu Ala Pro Ser Thr Thr Pro Ser Ser Ser Ala Tyr Ala
 450 455 460
 Thr Leu Ala Pro Thr Asp Arg Pro Pro Ser Arg Ser Ile Asp Phe Glu
 465 470 475 480
 Asp Leu Thr Ser Met Asp Thr Arg Ser Phe Ser Ser Asp Tyr Thr His
 485 490 495
 Leu Pro Glu Cys Gln Asn Pro Trp Asp Thr Asp Pro Pro Thr Tyr His
 500 505 510
 Thr Ile Glu Arg Ser Lys Ser Ser Arg Tyr Leu Ala Thr Thr Pro Phe
 515 520 525
 Leu Leu Glu Glu Ala Pro Ile Val Lys Ser His Ser Phe Met Phe Ser
 530 535 540
 Pro Ser Arg Ser Tyr Tyr Ala Asn Phe Gly Val Pro Val Lys Thr Ala
 545 550 555 560
 Glu Tyr Thr Ser Ile Thr Asp Cys Ile Asp Thr Arg Cys Val Asn Ala
 565 570 575
 Pro Gln Ala Ile Ala Asp Arg Ala Thr Phe Pro Gly Gly Leu Gly Asp
 580 585 590
 Lys Val Glu Asp Leu Ser Cys Cys His Pro Glu Arg Glu Ala Glu Leu
 595 600 605
 Ser His Pro Ser Ser Asp Ser Glu Glu Asn Glu Ala Arg Gly Gln Arg
 610 615 620
 Ala Ala Asn Pro Ile Ser Ser Gln Glu Ala Glu Asn Ala Asp Arg Thr
 625 630 635 640
 Leu Ser Asn Asn Ile Thr Val Pro Lys Ile Glu Arg Ala Asn Ser Tyr
 645 650 655
 Ser Ala Glu Glu Pro Asn Val Pro Tyr Ala His Thr Arg Lys Ser Phe
 660 665 670
 Ser Ile Ser Asp Lys Leu Asp Arg Gln Arg Asn Thr Ala Ser Leu Arg
 675 680 685

16U 100 PCT.ST25

Asn Pro Phe Gln Arg Lys Thr Ile Leu Gln Tyr Thr Pro Asn Lys Leu
690 695 700

Tyr Pro Glu Cys Leu Leu Ser Ser Ser Thr Gly Ala Val Glu Leu Tyr
705 710 715 720

Asp Pro Ala Glu Ala Ile Leu Leu Ala Ala Phe Leu Asp Gly Gly Tyr
725 730 735

<210> 21
<211> 24
<212> DNA
<213> Homo sapiens

<400> 21
gcttttagcct ggaacagagt cgac 24

<210> 22
<211> 24
<212> DNA
<213> Homo sapiens

<400> 22
gtcttttcttc ctgcctcaa ggga 24

<210> 23
<211> 23
<212> DNA
<213> Homo sapiens

<400> 23
gaccaaggaa tggcagttgg agc 23

<210> 24
<211> 24
<212> DNA
<213> Homo sapiens

<400> 24
gtggtcccgt tgtcagccag aatg 24

<210> 25
<211> 3456
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (342)..(1535)
<223>

<400> 25
ggacggtcca gaggtgtcga aatgtcctgg ggacctgagc agcagccacc agggaagagg 60
caggaggagg gctgaggacc aggcttggtt gtgagaatcc ctgagcccag gcggtagatg 120
ccaggagggtg tctggactgg ctgggccatg cctgggctga cctgtccagc caggagagg 180
gtgtgagggc agatctgggg gtgccagat ggaaggaggc aggcattggg gacaccaag 240
gccccctggc agcaccatga actaagcagg acacctggag gggaagaact gtggggacct 300
ggaggcctcc aacgactcct tcctgcttcc tggacaggac t atg gct gtg cag gga 356
Met Ala Val Gln Gly
1 5

tcc cag aga aga ctt ctg ggc tcc ctc aac tcc acc ccc aca gcc atc 404
Ser Gln Arg Arg Leu Leu Gly Ser Leu Asn Ser Thr Pro Thr Ala Ile
10 15 20

ccc cag ctg ggg ctg gct gcc aac cag aca gga gcc cgg tgc ctg gag 452
Pro Gln Leu Gly Leu Ala Ala Asn Gln Thr Gly Ala Arg Cys Leu Glu

25					30					160 100 PCT.ST25					35	
gtg	tcc	atc	tct	gac	ggg	ctc	ttc	ctc	agc	ctg	ggg	ctg	gtg	agc	ttg	500
Val	Ser	Ile	Ser	Asp	Gly	Leu	Phe	Leu	Ser	Leu	Gly	Leu	Val	Ser	Leu	
40					45					50						
gtg	gag	aac	gcg	ctg	gtg	gtg	gcc	acc	atc	gcc	aag	aac	cgg	aac	ctg	548
Val	Glu	Asn	Ala	Leu	Val	Val	Ala	Thr	Ile	Ala	Lys	Asn	Arg	Asn	Leu	
55					60					65						
cac	tca	ccc	atg	tac	tgc	ttc	atc	tgc	tgc	ctg	gcc	ttg	tcg	gac	ctg	596
His	Ser	Pro	Met	Tyr	Cys	Phe	Ile	Cys	Cys	Leu	Ala	Leu	Ser	Asp	Leu	
70					75					80					85	
ctg	gtg	agc	ggg	agc	aac	gtg	ctg	gag	acg	gcc	gtc	atc	ctc	ctg	ctg	644
Leu	Val	Ser	Gly	Ser	Asn	Val	Leu	Glu	Thr	Ala	Val	Ile	Leu	Leu	Leu	
90					95					100						
gag	gcc	ggt	gca	ctg	gtg	gcc	cgg	gct	gcg	gtg	ctg	cag	cag	ctg	gac	692
Glu	Ala	Gly	Ala	Leu	Val	Ala	Arg	Ala	Ala	Val	Leu	Gln	Gln	Leu	Asp	
105					110					115						
aat	gtc	act	gac	gtg	atc	acc	tgc	agc	tcc	atg	ctg	tcc	agc	ctc	tgc	740
Asn	Val	Thr	Asp	Val	Ile	Thr	Cys	Ser	Ser	Met	Leu	Ser	Ser	Leu	Cys	
120					125					130						
ttc	ctg	ggc	gcc	atc	gcc	gtg	gac	cgc	tac	atc	tcc	atc	ttc	tac	gca	788
Phe	Leu	Gly	Ala	Ile	Ala	Val	Asp	Arg	Tyr	Ile	Ser	Ile	Phe	Tyr	Ala	
135					140					145						
ctg	cgc	tac	cac	agc	atc	gtg	acc	ctg	ccg	cgg	gcg	cgg	cga	gcc	gtt	836
Leu	Arg	Tyr	His	Ser	Ile	Val	Thr	Leu	Pro	Arg	Ala	Arg	Arg	Ala	Val	
150					155					160					165	
gcg	gcc	atc	tgg	gtg	gcc	agt	gtc	gtc	ttc	agc	acg	ctc	ttc	atc	gcc	884
Ala	Ala	Ile	Trp	Val	Ala	Ser	Val	Val	Phe	Ser	Thr	Leu	Phe	Ile	Ala	
170					175					180						
tac	tac	gac	cac	gtg	gcc	gtc	ctg	ctg	tgc	ctc	gtg	gtc	ttc	ttc	ctg	932
Tyr	Tyr	Asp	His	Val	Ala	Val	Leu	Leu	Cys	Leu	Val	Val	Phe	Phe	Leu	
185					190					195						
gct	atg	ctg	gtg	ctc	atg	gcc	gtg	ctg	tac	gtc	cac	atg	ctg	gcc	cgg	980
Ala	Met	Leu	Val	Leu	Met	Ala	Val	Leu	Tyr	Val	His	Met	Leu	Ala	Arg	
200					205					210						
gcc	tgc	cag	cac	gcc	cag	ggc	atc	gcc	cgg	ctc	cac	aag	agg	cag	cgc	1028
Ala	Cys	Gln	His	Ala	Gln	Gly	Ile	Ala	Arg	Leu	His	Lys	Arg	Gln	Arg	
215					220					225						
ccg	gtc	cac	cag	ggc	ttt	ggc	ctt	aaa	ggc	gct	gtc	acc	ctc	acc	atc	1076
Pro	Val	His	Gln	Gly	Phe	Gly	Leu	Lys	Gly	Ala	Val	Thr	Leu	Thr	Ile	
230					235					240					245	
ctg	ctg	ggc	att	ttc	ttc	ctc	tgc	tgg	ggc	ccc	ttc	ttc	ctg	cat	ctc	1124
Leu	Leu	Gly	Ile	Phe	Phe	Leu	Cys	Trp	Gly	Pro	Phe	Phe	Leu	His	Leu	
250					255					260						
aca	ctc	atc	gtc	ctc	tgc	ccc	gag	cac	ccc	acg	tgc	ggc	tgc	atc	ttc	1172
Thr	Leu	Ile	Val	Leu	Cys	Pro	Glu	His	Pro	Thr	Cys	Gly	Cys	Ile	Phe	
265					270					275						
aag	aac	ttc	aac	ctc	ttt	ctc	gcc	ctc	atc	atc	tgc	aat	gcc	atc	atc	1220
Lys	Asn	Phe	Asn	Leu	Phe	Leu	Ala	Leu	Ile	Ile	Cys	Asn	Ala	Ile	Ile	
280					285					290						
gac	ccc	ctc	atc	tac	gcc	ttc	cac	agc	cag	gag	ctc	cgc	agg	acg	ctc	1268
Asp	Pro	Leu	Ile	Tyr	Ala	Phe	His	Ser	Gln	Glu	Leu	Arg	Arg	Thr	Leu	
295					300					305						
aag	gag	gtg	ctg	aca	tgc	tcc	tgc	tct	cag	gac	cgt	gcc	ctc	gtc	agc	1316
Lys	Glu	Val	Leu	Thr	Cys	Ser	Cys	Ser	Gln	Asp	Arg	Ala	Leu	Val	Ser	
310					315					320					325	
tgg	gat	gtg	aag	tct	ctg	ggg	gga	agt	gtg	tgc	caa	gag	cta	ctc	cca	1364
Trp	Asp	Val	Lys	Ser	Leu	Gly	Gly	Ser	Val	Cys	Gln	Glu	Leu	Leu	Pro	
330					335					340						
cag	cag	ccc	cag	gag	aag	ggg	ctt	tgt	gac	cag	aaa	gct	tca	tcc	aca	1412

160 100 PCT.ST25

Gln Gln Pro Gln Glu Lys Gly Leu Cys Asp Gln Lys Ala Ser Ser Thr	
345 350 355	
gcc ttg cag cgg ctc ctg caa aag gag cct aga gga agg acg agc agg	1460
Ala Leu Gln Arg Leu Leu Gln Lys Glu Pro Arg Gly Arg Thr Ser Arg	
360 365 370	
tgc agc agg gcc cca gtc ccc tcc act ctt gac gct gtc cta gct gca	1508
Cys Ser Arg Ala Pro Val Gln Pro Ser Thr Leu Asp Ala Val Leu Ala Ala	
375 380 385	
gaa gag gcg ggt tcc cag cct tcc ctg tgaccacatg tgacctcagc	1555
Glu Glu Ala Gly Ser Gln Pro Ser Leu	
390 395	
cgggacacat ccctttgctg gccctggccc tgagtccttc cagccatgat gagccgtgaa	1615
tgggaccatc cctgtccact ctgagatgcc tggagggggg ctcaagtgcag agactgagca	1675
ctcagtcagc ccccttcctg ggacaggctc aatggaggct gcagggccat cagccgactc	1735
ctacgcaggc tcagtcagca gcccctggc cagccccacc cctgactgcc ggccctcagaa	1795
ctgggagctg cttcctggca gggcccgct ctgctgggag accggacgtt ctgggaagtc	1855
atcagtgatg agcatggcat cgacccagc ggcaactacg tgggcgactc ggacttgca	1915
ctggagcgga tcagcgtcta ctacaacgag gcctcttctc acaagtacgt gcctcgagcc	1975
attctgtggt acctggaacc cggaaccatg gacagtgtcc gctcaggggc ctttgacat	2035
ctcttcaggc ctgacaattt catcttttgt cagagtgggg ccggcaacaa ctgggccaag	2095
ggtcactaca cggagggggc ggagctggtg gattcgggtc tggatgtggt gcggaaggag	2155
tgtgaaaact gcgactgcct gcagggcttc cagctgacct actcgtggg gggcggcacg	2215
ggctccggca tgggcacgtt gctcatcagc aaggtgcgtg aggagtatcc cgaccgcatc	2275
atgaacacct tcagcgtcgt gccctcacc aaggtgtcag acacggtggt ggagccctac	2335
aacgccacgc tgtccatcca ccagctggtg gagaacacgg atgagaccta ctgcatcgac	2395
aacgaggcgc tctacgacat ctgcttccgc accctcaagc tggccacgcc cacctacggg	2455
gacctcaacc acctggtatc gggcaccatg agcggagtca ccacctcctt gcgcttcccg	2515
ggccagctca acgctgacct gcgcaagctg gccgtcaaca tgggtgccctt cccgcgcctg	2575
cactttctca tgcccggtt cgccccctc acagcccggg gcagccagca gtaccggggc	2635
ctgaccgtgc ccgagctcac ccagcagatg ttcgatgcca agaacatgat ggccgcctgc	2695
gacccgcgcc acggccgcta cctgacggtg gccaccgtgt tccggggccg catgtccatg	2755
aaggaggtgg acgagcagat gctggccatc cagagcaaga acagcagcta cttcgtggag	2815
tggatcccca acaacgtgaa ggtggccgtg tgtgacatcc cgccccgcg cctcaagatg	2875
tcctccacct tcatcgggaa cagcacggcc atccaggagc tgttcaagcg catctccgag	2935
cagttcacgg ccattgtccg gcgcaaggcc ttcctgcact ggtacacggg cgagggcatg	2995
gacgagatgg agttcaccga ggccgagagc aacatgaacg acctggtgtc cgagtaccag	3055
cagtaccagg acgccacggc cgaggaagag ggcgagatgt acgaagacga cgaggaggag	3115
tcggaggccc agggcccaaa gtgaagctgc tcgcagctgg agtgagaggc aggtggcggc	3175
cggggccgaa gccagcagtg tctaaacccc cgagaccatc ttgctgccga caccctgctt	3235
tcccctcgcc ctagggtctc cttgccgccc tctgcagta tttatggcct cgtcctcccc	3295
acctaggcca cgtgtgagct gctcctgtct ctgtcttatt gcagctccag gcctgacgtt	3355
ttacggtttt gttttttact ggtttgtgtt tatattttcg gggatactta ataaatctat	3415
tgctgtcaga taaaaaaaa aaaaaaaaa aaaaaaaaa a	3456

16U 100 PCT.ST25

<210> 26
 <211> 398
 <212> PRT
 <213> Homo sapiens

<400> 26

```

Met Ala Val Gln Gly Ser Gln Arg Arg Leu Leu Gly Ser Leu Asn Ser
1      5      10      15

Thr Pro Thr Ala Ile Pro Gln Leu Gly Leu Ala Ala Asn Gln Thr Gly
      20      25      30

Ala Arg Cys Leu Glu Val Ser Ile Ser Asp Gly Leu Phe Leu Ser Leu
      35      40      45

Gly Leu Val Ser Leu Val Glu Asn Ala Leu Val Val Ala Thr Ile Ala
50      55      60

Lys Asn Arg Asn Leu His Ser Pro Met Tyr Cys Phe Ile Cys Cys Leu
65      70      75      80

Ala Leu Ser Asp Leu Leu Val Ser Gly Ser Asn Val Leu Glu Thr Ala
      85      90      95

Val Ile Leu Leu Leu Glu Ala Gly Ala Leu Val Ala Arg Ala Ala Val
100      105      110

Leu Gln Gln Leu Asp Asn Val Thr Asp Val Ile Thr Cys Ser Ser Met
115      120      125

Leu Ser Ser Leu Cys Phe Leu Gly Ala Ile Ala Val Asp Arg Tyr Ile
130      135      140

Ser Ile Phe Tyr Ala Leu Arg Tyr His Ser Ile Val Thr Leu Pro Arg
145      150      155      160

Ala Arg Arg Ala Val Ala Ala Ile Trp Val Ala Ser Val Val Phe Ser
165      170      175

Thr Leu Phe Ile Ala Tyr Tyr Asp His Val Ala Val Leu Leu Cys Leu
180      185      190

Val Val Phe Phe Leu Ala Met Leu Val Leu Met Ala Val Leu Tyr Val
195      200      205

His Met Leu Ala Arg Ala Cys Gln His Ala Gln Gly Ile Ala Arg Leu
210      215      220

His Lys Arg Gln Arg Pro Val His Gln Gly Phe Gly Leu Lys Gly Ala
225      230      235      240

Val Thr Leu Thr Ile Leu Leu Gly Ile Phe Phe Leu Cys Trp Gly Pro
245      250      255

Phe Phe Leu His Leu Thr Leu Ile Val Leu Cys Pro Glu His Pro Thr
260      265      270

Cys Gly Cys Ile Phe Lys Asn Phe Asn Leu Phe Leu Ala Leu Ile Ile
275      280      285

```


16U 100 PCT.ST25

Cys Asn Ala Ile Ile Asp Pro Leu Ile Tyr Ala Phe His Ser Gln Glu
 290 295 300

Leu Arg Arg Thr Leu Lys Glu Val Leu Thr Cys Ser Cys Ser Gln Asp
 305 310 315 320

Arg Ala Leu Val Ser Trp Asp Val Lys Ser Leu Gly Gly Ser Val Cys
 325 330 335

Gln Glu Leu Leu Pro Gln Gln Pro Gln Glu Lys Gly Leu Cys Asp Gln
 340 345 350

Lys Ala Ser Ser Thr Ala Leu Gln Arg Leu Leu Gln Lys Glu Pro Arg
 355 360 365

Gly Arg Thr Ser Arg Cys Ser Arg Ala Pro Val Pro Ser Thr Leu Asp
 370 375 380

Ala Val Leu Ala Ala Glu Glu Ala Gly Ser Gln Pro Ser Leu
 385 390 395

<210> 27
 <211> 398
 <212> PRT
 <213> Homo sapiens

<400> 27

Met Ala Val Gln Gly Ser Gln Arg Arg Leu Leu Gly Ser Leu Asn Ser
 1 5 10 15

Thr Pro Thr Ala Ile Pro Gln Leu Gly Leu Ala Ala Asn Gln Thr Gly
 20 25 30

Ala Arg Cys Leu Glu Val Ser Ile Ser Asp Gly Leu Phe Leu Ser Leu
 35 40 45

Gly Leu Val Ser Leu Val Glu Asn Ala Leu Val Val Ala Thr Ile Ala
 50 55 60

Lys Asn Arg Asn Leu His Ser Pro Met Tyr Cys Phe Ile Cys Cys Leu
 65 70 75 80

Ala Leu Ser Asp Leu Leu Val Ser Gly Ser Asn Val Leu Glu Thr Ala
 85 90 95

Val Ile Leu Leu Leu Glu Ala Gly Ala Leu Val Ala Arg Ala Ala Val
 100 105 110

Leu Gln Gln Leu Asp Asn Val Thr Asp Val Ile Thr Cys Ser Ser Met
 115 120 125

Leu Ser Ser Leu Cys Phe Leu Gly Ala Ile Ala Val Asp Arg Tyr Ile
 130 135 140

Ser Ile Phe Tyr Ala Leu Arg Tyr His Ser Ile Val Thr Leu Pro Arg
 145 150 155 160

Ala Arg Gln Ala Val Ala Ala Ile Trp Val Ala Ser Val Val Phe Ser
 165 170 175

160 100 PCT.ST25

Thr Leu Phe Ile Ala Tyr Tyr Asp His Val Ala Val Leu Leu Cys Leu
180 185 190

Val Val Phe Phe Leu Ala Met Leu Val Leu Met Ala Val Leu Tyr Val
195 200 205

His Met Leu Ala Arg Ala Cys Gln His Ala Gln Gly Ile Ala Arg Leu
210 215 220

His Lys Arg Gln Arg Pro Val His Gln Gly Phe Gly Leu Lys Gly Ala
225 230 235 240

Val Thr Leu Thr Ile Leu Leu Gly Ile Phe Phe Leu Cys Trp Gly Pro
245 250 255

Phe Phe Leu His Leu Thr Leu Ile Val Leu Cys Pro Glu His Pro Thr
260 265 270

Cys Gly Cys Ile Phe Lys Asn Phe Asn Leu Phe Leu Ala Leu Ile Ile
275 280 285

Cys Asn Ala Ile Ile Asp Pro Leu Ile Tyr Ala Phe His Ser Gln Glu
290 295 300

Leu Arg Arg Thr Leu Lys Glu Val Leu Thr Cys Ser Cys Ser Gln Asp
305 310 315 320

Arg Ala Leu Val Ser Trp Asp Val Lys Ser Leu Gly Gly Ser Val Cys
325 330 335

Gln Glu Leu Leu Pro Gln Gln Pro Gln Glu Lys Gly Leu Cys Asp Gln
340 345 350

Lys Ala Ser Ser Thr Ala Leu Gln Arg Leu Leu Gln Lys Glu Pro Arg
355 360 365

Gly Arg Thr Ser Arg Cys Ser Arg Ala Pro Val Pro Ser Thr Leu Asp
370 375 380

Ala Val Leu Ala Ala Glu Glu Ala Gly Ser Gln Pro Ser Leu
385 390 395

<210> 28
<211> 398
<212> PRT
<213> Homo sapiens

<400> 28

Met Ala Val Gln Gly Ser Gln Arg Arg Leu Leu Gly Ser Leu Asn Ser
1 5 10 15

Thr Pro Thr Ala Ile Pro Gln Leu Gly Leu Ala Ala Asn Gln Thr Gly
20 25 30

Ala Arg Cys Leu Glu Val Ser Ile Ser Asp Gly Leu Phe Leu Ser Leu
35 40 45

Gly Leu Val Ser Leu Val Glu Asn Ala Leu Val Val Ala Thr Ile Ala
50 55 60

16U 100 PCT.ST25

Lys Asn Arg Asn Leu His Ser Pro Met Tyr Cys Phe Ile Cys Cys Leu
 65 70 75 80
 Ala Leu Ser Asp Leu Leu Val Ser Gly Ser Asn Val Leu Glu Thr Ala
 85 90 95
 Val Ile Leu Leu Leu Glu Ala Gly Ala Leu Val Ala Arg Ala Ala Val
 100 105 110
 Leu Gln Gln Leu Asp Asn Val Ile Asp Val Ile Thr Cys Ser Ser Met
 115 120 125
 Leu Ser Ser Leu Cys Phe Leu Gly Ala Ile Ala Val Asp Arg Tyr Ile
 130 135 140
 Ser Ile Phe Tyr Ala Leu Arg Tyr His Ser Ile Val Thr Leu Pro Arg
 145 150 155 160
 Ala Arg Arg Ala Val Ala Ala Ile Trp Val Ala Ser Val Val Phe Ser
 165 170 175
 Thr Leu Phe Ile Ala Tyr Tyr Asp His Val Ala Val Leu Leu Cys Leu
 180 185 190
 Val Val Phe Phe Leu Ala Met Leu Val Leu Met Ala Val Leu Tyr Val
 195 200 205
 His Met Leu Ala Arg Ala Cys Gln His Ala Gln Gly Ile Ala Arg Leu
 210 215 220
 His Lys Arg Gln Arg Pro Val His Gln Gly Phe Gly Leu Lys Gly Ala
 225 230 235 240
 Val Thr Leu Thr Ile Leu Leu Gly Ile Phe Phe Leu Cys Trp Gly Pro
 245 250 255
 Phe Phe Leu His Leu Thr Leu Ile Val Leu Cys Pro Glu His Pro Thr
 260 265 270
 Cys Gly Cys Ile Phe Lys Asn Phe Asn Leu Phe Leu Ala Leu Ile Ile
 275 280 285
 Cys Asn Ala Ile Ile Asp Pro Leu Ile Tyr Ala Phe His Ser Gln Glu
 290 295 300
 Leu Arg Arg Thr Leu Lys Glu Val Leu Thr Cys Ser Cys Ser Gln Asp
 305 310 315 320
 Arg Ala Leu Val Ser Trp Asp Val Lys Ser Leu Gly Gly Ser Val Cys
 325 330 335
 Gln Glu Leu Leu Pro Gln Gln Pro Gln Glu Lys Gly Leu Cys Asp Gln
 340 345 350
 Lys Ala Ser Ser Thr Ala Leu Gln Arg Leu Leu Gln Lys Glu Pro Arg
 355 360 365
 Gly Arg Thr Ser Arg Cys Ser Arg Ala Pro Val Pro Ser Thr Leu Asp
 370 375 380

16U 100 PCT.ST25

Ala Val Leu Ala Ala Glu Glu Ala Gly Ser Gln Pro Ser Leu
 385 390 395

<210> 29
 <211> 398
 <212> PRT
 <213> Homo sapiens

<400> 29

Met Ala Val Gln Gly Ser Gln Arg Arg Leu Leu Gly Ser Leu Asn Ser
 1 5 10 15

Thr Pro Thr Ala Ile Pro Gln Leu Gly Leu Ala Ala Asn Gln Thr Gly
 20 25 30

Ala Arg Cys Leu Glu Val Ser Ile Ser Asp Gly Leu Phe Leu Ser Leu
 35 40 45

Gly Leu Val Ser Leu Val Glu Asn Ala Leu Val Val Ala Thr Ile Ala
 50 55 60

Lys Asn Arg Asn Leu His Ser Pro Met Tyr Cys Phe Ile Cys Cys Leu
 65 70 75 80

Ala Leu Ser Asp Leu Leu Val Ser Gly Ser Asn Val Leu Glu Thr Ala
 85 90 95

Val Ile Leu Leu Leu Glu Ala Gly Ala Leu Val Ala Arg Ala Ala Val
 100 105 110

Leu Gln Gln Leu Asp Asn Val Ile Asp Val Ile Thr Cys Ser Ser Met
 115 120 125

Leu Ser Ser Leu Cys Phe Leu Gly Ala Ile Ala Val Asp Arg Tyr Ile
 130 135 140

Ser Ile Phe Tyr Ala Leu Arg Tyr His Ser Ile Val Thr Leu Pro Arg
 145 150 155 160

Ala Arg Gln Ala Val Ala Ala Ile Trp Val Ala Ser Val Val Phe Ser
 165 170 175

Thr Leu Phe Ile Ala Tyr Tyr Asp His Val Ala Val Leu Leu Cys Leu
 180 185 190

Val Val Phe Phe Leu Ala Met Leu Val Leu Met Ala Val Leu Tyr Val
 195 200 205

His Met Leu Ala Arg Ala Cys Gln His Ala Gln Gly Ile Ala Arg Leu
 210 215 220

His Lys Arg Gln Arg Pro Val His Gln Gly Phe Gly Leu Lys Gly Ala
 225 230 235 240

Val Thr Leu Thr Ile Leu Leu Gly Ile Phe Phe Leu Cys Trp Gly Pro
 245 250 255

Phe Phe Leu His Leu Thr Leu Ile Val Leu Cys Pro Glu His Pro Thr
 260 265 270

16U 100 PCT.ST25

Cys Gly Cys Ile Phe Lys Asn Phe Asn Leu Phe Leu Ala Leu Ile Ile
 275 280 285

Cys Asn Ala Ile Ile Asp Pro Leu Ile Tyr Ala Phe His Ser Gln Glu
 290 295 300

Leu Arg Arg Thr Leu Lys Glu Val Leu Thr Cys Ser Cys Ser Gln Asp
 305 310 315 320

Arg Ala Leu Val Ser Trp Asp Val Lys Ser Leu Gly Gly Ser Val Cys
 325 330 335

Gln Glu Leu Leu Pro Gln Gln Pro Gln Glu Lys Gly Leu Cys Asp Gln
 340 345 350

Lys Ala Ser Ser Thr Ala Leu Gln Arg Leu Leu Gln Lys Glu Pro Arg
 355 360 365

Gly Arg Thr Ser Arg Cys Ser Arg Ala Pro Val Pro Ser Thr Leu Asp
 370 375 380

Ala Val Leu Ala Ala Glu Glu Ala Gly Ser Gln Pro Ser Leu
 385 390 395

<210> 30
 <211> 317
 <212> PRT
 <213> Homo sapiens

<400> 30

Met Ala Val Gln Gly Ser Gln Arg Arg Leu Leu Gly Ser Leu Asn Ser
 1 5 10 15

Thr Pro Thr Ala Ile Pro Gln Leu Gly Leu Ala Ala Asn Gln Thr Gly
 20 25 30

Ala Arg Cys Leu Glu Val Ser Ile Ser Asp Gly Leu Phe Leu Ser Leu
 35 40 45

Gly Leu Val Ser Leu Val Glu Asn Ala Leu Val Val Ala Thr Ile Ala
 50 55 60

Lys Asn Arg Asn Leu His Ser Pro Met Tyr Cys Phe Ile Cys Cys Leu
 65 70 75 80

Ala Leu Ser Asp Leu Leu Val Ser Gly Ser Asn Val Leu Glu Thr Ala
 85 90 95

Val Ile Leu Leu Leu Glu Ala Gly Ala Leu Val Ala Arg Ala Ala Val
 100 105 110

Leu Gln Gln Leu Asp Asn Val Ile Asp Val Ile Thr Cys Ser Ser Met
 115 120 125

Leu Ser Ser Leu Cys Phe Leu Gly Ala Ile Ala Val Asp Arg Tyr Ile
 130 135 140

Ser Ile Phe Tyr Ala Leu Arg Tyr His Ser Ile Val Thr Leu Pro Arg
 145 150 155 160

16U 100 PCT.ST25

Ala Arg Gln Ala Val Ala Ala Ile Trp Val Ala Ser Val Val Phe Ser
165 170 175

Thr Leu Phe Ile Ala Tyr Tyr Asp His Val Ala Val Leu Leu Cys Leu
180 185 190

Val Val Phe Phe Leu Ala Met Leu Val Leu Met Ala Val Leu Tyr Val
195 200 205

His Met Leu Ala Arg Ala Cys Gln His Ala Gln Gly Ile Ala Arg Leu
210 215 220

His Lys Arg Gln Arg Pro Val His Gln Gly Phe Gly Leu Lys Gly Ala
225 230 235 240

Val Thr Leu Thr Ile Leu Leu Gly Ile Phe Phe Leu Cys Trp Gly Pro
245 250 255

Phe Phe Leu His Leu Thr Leu Ile Val Leu Cys Pro Glu His Pro Thr
260 265 270

Cys Gly Cys Ile Phe Lys Asn Phe Asn Leu Phe Leu Ala Leu Ile Ile
275 280 285

Cys Asn Ala Ile Ile Asp Pro Leu Ile Tyr Ala Phe His Ser Gln Glu
290 295 300

Leu Arg Arg Thr Leu Lys Glu Val Leu Thr Cys Ser Trp
305 310 315

<210> 31
<211> 382
<212> PRT
<213> Homo sapiens

<400> 31

Met Ala Val Gln Gly Ser Gln Arg Arg Leu Leu Gly Ser Leu Asn Ser
1 5 10 15

Thr Pro Thr Ala Ile Pro Gln Leu Gly Leu Ala Ala Asn Gln Thr Gly
20 25 30

Ala Arg Cys Leu Glu Val Ser Ile Ser Asp Gly Leu Phe Leu Ser Leu
35 40 45

Gly Leu Val Ser Leu Val Glu Asn Ala Leu Val Val Ala Thr Ile Ala
50 55 60

Lys Asn Arg Asn Leu His Ser Pro Met Tyr Cys Phe Ile Cys Cys Leu
65 70 75 80

Ala Leu Ser Asp Leu Leu Val Ser Gly Ser Asn Val Leu Glu Thr Ala
85 90 95

Val Ile Leu Leu Leu Glu Ala Gly Ala Leu Val Ala Arg Ala Ala Val
100 105 110

Leu Gln Gln Leu Asp Asn Val Ile Asp Val Ile Thr Cys Ser Ser Met
115 120 125

160 100 PCT.ST25

Leu Ser Ser Leu Cys Phe Leu Gly Ala Ile Ala Val Asp Arg Tyr Ile
 130 135 140

Ser Ile Phe Tyr Ala Leu Arg Tyr His Ser Ile Val Thr Leu Pro Arg
 145 150 155 160

Ala Arg Arg Ala Val Ala Ala Ile Trp Val Ala Ser Val Val Phe Ser
 165 170 175

Thr Leu Phe Ile Ala Tyr Tyr Asp His Val Ala Val Leu Leu Cys Leu
 180 185 190

Val Val Phe Phe Leu Ala Met Leu Val Leu Met Ala Val Leu Tyr Val
 195 200 205

His Met Leu Ala Arg Ala Cys Gln His Ala Gln Gly Ile Ala Arg Leu
 210 215 220

His Lys Arg Gln Arg Pro Val His Gln Gly Phe Gly Leu Lys Gly Ala
 225 230 235 240

Val Thr Leu Thr Ile Leu Leu Gly Ile Phe Phe Leu Cys Trp Gly Pro
 245 250 255

Phe Phe Leu His Leu Thr Leu Ile Val Leu Cys Pro Glu His Pro Thr
 260 265 270

Cys Gly Cys Ile Phe Lys Asn Phe Asn Leu Phe Leu Ala Leu Ile Ile
 275 280 285

Cys Asn Ala Ile Ile Asp Pro Leu Ile Tyr Ala Phe His Ser Gln Glu
 290 295 300

Leu Arg Arg Thr Leu Lys Glu Val Leu Thr Cys Ser Cys Ser Gln Asp
 305 310 315 320

Arg Ala Leu Val Ser Trp Asp Val Lys Ser Leu Gly Gly Ser Val Cys
 325 330 335

Gln Glu Leu Leu Pro Gln Gln Pro Gln Glu Lys Gly Leu Cys Asp Gln
 340 345 350

Lys Ala Ser Ser Thr Ala Leu Gln Arg Leu Leu Gln Lys Glu Val Lys
 355 360 365

Ser Leu Pro Gln Ala Lys Gly Pro Gly Leu Gln Glu Pro Pro
 370 375 380

<210> 32
 <211> 22
 <212> DNA
 <213> Homo sapiens

<400> 32
 cctcatcatc tgcaatgccca tc

22

<210> 33
 <211> 22
 <212> DNA
 <213> Homo sapiens

16U 100 PCT.ST25

<400> 33
 gctcgtcctt cctctaggct cc 22

<210> 34
 <211> 22
 <212> DNA
 <213> Homo sapiens

<400> 34
 aggaagcagc tcccagttct ga 22

<210> 35
 <211> 50
 <212> DNA
 <213> Homo sapiens

<400> 35
 cttccgcagc ggaaatggcg cgccgcccgg ggagggcggg agcagcgtcc 50

<210> 36
 <211> 50
 <212> DNA
 <213> Homo sapiens

<400> 36
 cctcaggctc tacaagatgc ctgaaaacac caacctctcc agggctcact 50

<210> 37
 <211> 50
 <212> DNA
 <213> Homo sapiens

<400> 37
 aacgactttt taaaacgcag agaaaagctc cattcttccc aggacctcag 50

<210> 38
 <211> 7062
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (186)..(5288)
 <223>

<400> 38
 cggtcgcagc gcacagagct tgctggccag ggaggagcta gtctccgtgg gcgcccgcgc 60
 cgccccagcc tgcgcgcctc tctcctggcg cgcgccagtc tggcactctg ggagctgggt 120
 cctagcacca cagacttata cttcgcctgc actttccgtc tttcttctct gggcgccac 180
 caaca atg gat ggc aac tcc ctg ctc tcg gta cca agc aac ttg gag tca 230
 Met Asp Gly Asn Ser Leu Leu Ser Val Pro Ser Asn Leu Glu Ser
 1 5 10 15
 tca cgg atg tat gac gtt ttg gaa ccg cag cag ggc aga ggc tgt ggc 278
 Ser Arg Met Tyr Asp Val Leu Glu Pro Gln Gln Gly Arg Gly Cys Gly
 20 25 30
 agc tca gga agc ggc ccg ggg aac tcc atc aca gcc tgt aag aag gtt 326
 Ser Ser Gly Ser Gly Pro Gly Asn Ser Ile Thr Ala Cys Lys Lys Val
 35 40 45
 ctt cgc agc aat agc ctg ctg gag tca aca gac tac tgg ttg cag aat 374
 Leu Arg Ser Asn Ser Leu Leu Glu Ser Thr Asp Tyr Trp Leu Gln Asn
 50 55 60
 cag agg atg ccc tgc caa att ggt ttt gta gaa gac aag tct gaa aac 422
 Gln Arg Met Pro Cys Gln Ile Gly Phe Val Glu Asp Lys Ser Glu Asn
 65 70 75
 tgt gct tct gtc tgc ttt gtg aat ctt gat gtg aac aag gat gaa tgc 470

160 100 PCT.ST25

Cys 80	Ala	Ser	Val	Cys	Phe 85	Val	Asn	Leu	Asp	Val 90	Asn	Lys	Asp	Glu	Cys 95	
agc Ser	aca Thr	gag Glu	cac His	ctg Leu	caa Gln	cag Gln	aaa Lys	ctg Leu	gtc Val	aac Asn	gtt Val	tca Ser	cca Pro	gat Asp	ctt Leu	518
			100					105						110		
cca Pro	aaa Lys	ctt Leu	atc Ile	agt Ser	tcc Ser	atg Met	aat Asn	gtc Val	caa Gln	caa Gln	cca Pro	aaa Lys	gaa Glu	aat Asn	gaa Glu	566
			115					120					125			
att Ile	gtt Val	gtc Val	cta Leu	agt Ser	ggg Gly	tta Leu	gcc Ala	tct Ser	gga Gly	aat Asn	ctc Leu	cag Gln	gca Ala	gat Asp	ttt Phe	614
			130				135					140				
gaa Glu	gtt Val	tca Ser	cag Gln	tgc Cys	cct Pro	tgg Trp	ctg Leu	cca Pro	gat Asp	atc Ile	tgc Cys	ttg Leu	gtc Val	caa Gln	tgt Cys	662
			145			150					155					
gca Ala	aga Arg	ggg Gly	aac Asn	aga Arg	cca Pro	aac Asn	agt Ser	acc Thr	aac Asn	tgc Cys	atc Ile	atc Ile	ttt Phe	gaa Glu	atc Ile	710
					160	165				170					175	
aac Asn	aaa Lys	ttt Phe	ctg Leu	att Ile	ggg Gly	ctg Leu	gaa Glu	ctg Leu	gtg Val	cag Gln	gag Glu	cga Arg	cag Gln	ctc Leu	cac His	758
			180						185					190		
ctg Leu	gaa Glu	aca Thr	aac Asn	atc Ile	ttg Leu	aaa Lys	ctg Leu	gag Glu	gat Asp	gac Asp	acg Thr	aac Asn	tgt Cys	tcc Ser	tta Leu	806
			195					200					205			
tct Ser	tca Ser	atc Ile	gag Glu	gaa Glu	gac Asp	ttt Phe	ctc Leu	acc Thr	gct Ala	tct Ser	gag Glu	cac His	ttg Leu	gag Glu	gag Glu	854
			210				215					220				
gaa Glu	agc Ser	gag Glu	gtg Val	gat Asp	gaa Glu	tct Ser	agg Arg	aac Asn	gat Asp	tat Tyr	gaa Glu	aat Asn	ata Ile	aat Asn	gtc Val	902
			225			230					235					
tca Ser	gcc Ala	aat Asn	gtt Val	ttg Leu	gaa Glu	agt Ser	aaa Lys	cag Gln	cta Leu	aag Lys	gga Gly	gcc Ala	acc Thr	cag Gln	gtg Val	950
					240	245				250				255		
gaa Glu	tgg Trp	aat Asn	tgc Cys	aac Asn	aag Lys	gaa Glu	aag Lys	tgg Trp	ctt Leu	tat Tyr	gct Ala	ttg Leu	gaa Glu	gac Asp	aaa Lys	998
				260					265					270		
tac Tyr	atc Ile	aac Asn	aaa Lys	tat Tyr	ccc Pro	aca Thr	cca Pro	ttg Leu	att Ile	aaa Lys	aca Thr	gaa Glu	cga Arg	tct Ser	cca Pro	1046
			275					280					285			
gaa Glu	aac Asn	cta Leu	aca Thr	aag Lys	aac Asn	aca Thr	gcc Ala	ttg Leu	cag Gln	agt Ser	cta Leu	gat Asp	ccc Pro	tca Ser	gcc Ala	1094
			290				295					300				
aag Lys	cca Pro	tca Ser	cag Gln	tgg Trp	aaa Lys	aga Arg	gaa Glu	gct Ala	gtg Val	ggg Gly	aat Asn	ggg Gly	aga Arg	caa Gln	gcc Ala	1142
			305			310					315					
aca Thr	cat His	tat Tyr	tat Tyr	cat His	tca Ser	gaa Glu	gct Ala	ttt Phe	aaa Lys	ggt Gly	caa Gln	atg Met	gaa Glu	aaa Lys	tca Ser	1190
					320	325			330					335		
cag Gln	gca Ala	ctg Leu	tat Tyr	att Ile	cca Pro	aaa Lys	gat Asp	gct Ala	tat Tyr	ttc Phe	tcc Ser	atg Met	atg Met	gat Asp	aaa Lys	1238
				340				345						350		
gat Asp	gta Val	cct Pro	tct Ser	gca Ala	tgt Cys	gct Ala	gtg Val	gca Ala	gag Glu	cag Gln	aga Arg	agc Ser	aac Asn	cta Leu	aac Asn	1286
			355				360						365			
cca Pro	gga Gly	gac Asp	cat His	gaa Glu	gac Asp	aca Thr	aga Arg	aac Asn	gct Ala	ctc Leu	cct Pro	cct Pro	aga Arg	caa Gln	gat Asp	1334
			370				375				380					
gga Gly	gaa Glu	gtc Val	acc Thr	act Thr	ggc Gly	aag Lys	tat Tyr	gct Ala	aca Thr	aat Asn	tta Leu	gca Ala	gaa Glu	tcc Ser	gtg Val	1382
			385			390					395					

16U 100 PCT.ST25

ctg cag gat gca ttt att aga tta tct caa tct cag tcc aca tta ccc Leu Gln Asp Ala Phe Ile Arg Leu Ser Gln Ser Gln Ser Thr Leu Pro 400 405 410 415	1430
cag gaa tct gca gtc agt gtt tct gta gga agt tct ctg ctt ccc agt Gln Glu Ser Ala Val Ser Val Ser Val Gly Ser Ser Leu Leu Pro Ser 420 425 430	1478
tgc tat tcc aca aaa gat aca gtg gtt tct cgg tca tgg aat gag ctc Cys Tyr Ser Thr Lys Asp Thr Val Ser Arg Ser Trp Asn Glu Leu 435 440 445	1526
ccc aaa atc gtc gtt gtt cag agt cca gat ggc agt gat gct gcc cca Pro Lys Ile Val Val Val Gln Ser Pro Asp Gly Ser Asp Ala Ala Pro 450 455 460	1574
cag cca ggc atc tcc tcc tgg cct gag atg gaa gtc tct gtt gaa acc Gln Pro Gly Ile Ser Ser Trp Pro Glu Met Glu Val Ser Val Glu Thr 465 470 475	1622
tca agc atc ctc tct gga gag aac tcc agc aga caa ccc cag agt gct Ser Ser Ile Leu Ser Gly Glu Asn Ser Ser Arg Gln Pro Gln Ser Ala 480 485 490 495	1670
cta gaa gtg gcg tta gct tgt gca gcc act gtg att gga act att tcc Leu Glu Val Ala Leu Ala Cys Ala Ala Thr Val Ile Gly Thr Ile Ser 500 505 510	1718
agt cca cag gcc aca gaa aga ctc aaa atg gag caa gtg gtc tcg aac Ser Pro Gln Ala Thr Glu Arg Leu Lys Met Glu Gln Val Val Ser Asn 515 520 525	1766
ttt ccc cca ggg agc agt ggt gca ctg caa act caa gca ccc caa gga Phe Pro Pro Gly Ser Ser Gly Ala Leu Gln Thr Gln Ala Pro Gln Gly 530 535 540	1814
ctc aag gaa cct tcc atc aat gag tac tcc ttt cca tct gct ttg tgt Leu Lys Glu Pro Ser Ile Asn Glu Tyr Ser Phe Pro Ser Ala Leu Cys 545 550 555	1862
ggc atg act cag gtg gcc agt gcc gtg gct gtc tgt ggt ctg ggt gaa Gly Met Thr Gln Val Ala Ser Ala Val Ala Val Cys Gly Leu Gly Glu 560 565 570 575	1910
aga gaa gag gtg aca tgc tca gtg gct cca agt ggt agc ctc ccg cct Arg Glu Glu Val Thr Cys Ser Val Ala Pro Ser Gly Ser Leu Pro Pro 580 585 590	1958
gca gct gag gct tct gaa gcc atg ccc cca ctt tgt ggt tta gca agc Ala Ala Glu Ala Ser Glu Ala Met Pro Pro Leu Cys Gly Leu Ala Ser 595 600 605	2006
atg gag ctt ggc aag gaa gcc att gcc gag gga ttg ctc aag gag gct Met Glu Leu Gly Lys Glu Ala Ile Ala Glu Gly Leu Leu Lys Glu Ala 610 615 620	2054
gct ctg gtt tta aca agg cct aat acc tac agc agc att gga gac ttt Ala Leu Val Leu Thr Arg Pro Asn Thr Tyr Ser Ser Ile Gly Asp Phe 625 630 635	2102
ctg gac tcc atg aac agg aga atc atg gaa act gct tca aag tct cag Leu Asp Ser Met Asn Arg Arg Ile Met Glu Thr Ala Ser Lys Ser Gln 640 645 650 655	2150
acc ctg tgc tca gaa aat gtc gtc agg aat gaa ctg gca cat acc ctg Thr Leu Cys Ser Glu Asn Val Val Arg Asn Glu Leu Ala His Thr Leu 660 665 670	2198
tcc aat gtt atc ctg agg cat tcc att gat gaa gtt cac cac aaa aat Ser Asn Val Ile Leu Arg His Ser Ile Asp Glu Val His His Lys Asn 675 680 685	2246
atg ata atc gac ccc aat gac aac agg cat tca tct gaa att ctg gac Met Ile Ile Asp Pro Asn Asp Asn Arg His Ser Ser Glu Ile Leu Asp 690 695 700	2294
acc tta atg gaa agt aca aat caa ctg ctt tta gat gtg ata tgc ttc Thr Leu Met Glu Ser Thr Asn Gln Leu Leu Leu Asp Val Ile Cys Phe 705 710 715	2342

160 100 PCT.ST25

acg ttc aag aag atg agt cat att gta cgg ctt ggt gaa tgt cct gct Thr Phe Lys Lys Met Ser His Ile Val Arg Leu Gly Glu Cys Pro Ala 720 725 730 735	2390
gtc ctt tct aag gag acc atc aga agg agg gag aca gaa cca agc tgc Val Leu Ser Lys Glu Thr Ile Arg Arg Glu Thr Glu Pro Ser Cys 740 745 750	2438
cag cca tct gat ccg ggt gct agt caa gct tgg aca aaa gcc act gaa Gln Pro Ser Asp Pro Gly Ala Ser Gln Ala Trp Thr Lys Ala Thr Glu 755 760 765	2486
tcc tcc agc agc tct cca ctt agc aat tca cac aac acg agt ctt gtc Ser Ser Ser Ser Ser Pro Leu Ser Asn Ser His Asn Thr Ser Leu Val 770 775 780	2534
atc aac aat ctt gtg gat ggc atg tat tca aaa caa gac aag ggt gga Ile Asn Asn Leu Val Asp Gly Met Tyr Ser Lys Gln Asp Lys Gly Gly 785 790 795	2582
gtg agg cca ggc ctc ttc aag aac ccc acg ctg cag tca caa tta tca Val Arg Pro Gly Leu Phe Lys Asn Pro Thr Leu Gln Ser Gln Leu Ser 800 805 810 815	2630
cgt agt cac aga gtg ccc gat tct tca act gct aca aca tcc tcc aag Arg Ser His Arg Val Pro Asp Ser Ser Thr Ala Thr Thr Ser Ser Lys 820 825 830	2678
gaa ata tat ctg aaa gga ata gca gga gag gat aca aaa agc cct cat Glu Ile Tyr Leu Lys Gly Ile Ala Gly Glu Asp Thr Lys Ser Pro His 835 840 845	2726
cac agt gag aat gaa tgc aga gcc tct tcc gaa gga caa agg tcc cca His Ser Glu Ala Glu Cys Arg Ala Ser Ser Glu Gly Gln Arg Ser Pro 850 855 860	2774
acg gtc agc cgg tcc aga agt ggt tcc cag gag gct gag gag agt atc Thr Val Ser Arg Ser Arg Ser Gly Ser Gln Glu Ala Glu Glu Ser Ile 865 870 875	2822
cac cca aac acc caa gaa aag tac aac tgt gcc aca tct cgc atc aac His Pro Asn Thr Gln Glu Lys Tyr Asn Cys Ala Thr Ser Arg Ile Asn 880 885 890 895	2870
gaa gtt caa gtc aac ctg tcc ttg tta ggg gat gac ctg ctg ctt cct Glu Val Gln Val Asn Leu Ser Leu Leu Gly Asp Asp Leu Leu Leu Pro 900 905 910	2918
gct caa tcc acg ctt caa aca aag cat cca gac atc tac tgc att aca Ala Gln Ser Thr Leu Gln Thr Lys His Pro Asp Ile Tyr Cys Ile Thr 915 920 925	2966
gac ttt gcg gaa gaa tta gca gac acg gtc gtc tcc atg gca act gaa Asp Phe Ala Glu Glu Leu Ala Asp Thr Val Val Ser Met Ala Thr Glu 930 935 940	3014
att gca gcg att tgc ctt gac aac tcc agt gga aaa caa ccc tgg ttt Ile Ala Ala Ile Cys Leu Asp Asn Ser Ser Gly Lys Gln Pro Trp Phe 945 950 955	3062
tgt gca tgg aaa aga ggg agt gag ttt ctg atg aca ccc aac gta ccc Cys Ala Trp Lys Arg Gly Ser Glu Phe Leu Met Thr Pro Asn Val Pro 960 965 970 975	3110
tgc cga tcc ttg aag agg aag aaa gag agc cag ggg agc ggg acc gct Cys Arg Ser Leu Lys Arg Lys Lys Glu Ser Gln Gly Ser Gly Thr Ala 980 985 990	3158
gtg agg aaa cac aag cct ccc cgg ctc agt gag atc aag agg aag acg Val Arg Lys His Lys Pro Pro Arg Leu Ser Glu Ile Lys Arg Lys Thr 995 1000 1005	3206
gac gag cac cct gag ctt aaa gaa aag ctg atg aac agg gtt gtg Asp Glu His Pro Glu Leu Lys Glu Lys Leu Met Asn Arg Val Val 1010 1015 1020	3251
gat gag tcc atg aac ctt gaa gat gtc cca gat tct gtc aat ctt Asp Glu Ser Met Asn Leu Glu Asp Val Pro Asp Ser Val Asn Leu	3296

1025	1030	16U 100 PCT.ST25 1035	
ttt gcc aat gaa gtg gca gcc aag atc atg aac cta acg gag ttc Phe Ala Asn Glu Val Ala Ala Lys Ile Met Asn Leu Thr Glu Phe 1040 1045 1050			3341
tct atg gtg gac ggc atg tgg cag gcg cag ggc tat ccc cgg aat Ser Met Val Asp Gly Met Trp Gln Ala Gln Gly Tyr Pro Arg Asn 1055 1060 1065			3386
cgg tta ctg agt ggc gac agg tgg agc cgg ctg aag gcc tcc agc Arg Leu Leu Ser Gly Asp Arg Trp Ser Arg Leu Lys Ala Ser Ser 1070 1075 1080			3431
tgc gaa agc att cct gag gaa gac tcc gag gcc agg gcc tat gtc Cys Glu Ser Ile Pro Glu Glu Asp Ser Glu Ala Arg Ala Tyr Val 1085 1090 1095			3476
aac agc ctg ggc tta atg agc acg ctg agc cag ccg gtc agc agg Asn Ser Leu Gly Leu Met Ser Thr Leu Ser Gln Pro Val Ser Arg 1100 1105 1110			3521
gcc agc tct gtc tcc aag cag tcc agc tgt gag agc atc acc gat Ala Ser Ser Val Ser Lys Gln Ser Ser Cys Glu Ser Ile Thr Asp 1115 1120 1125			3566
gag ttt tcc agg ttc atg gtg aac cag atg gaa aat gaa ggg aga Glu Phe Ser Arg Phe Met Val Asn Gln Met Glu Asn Glu Gly Arg 1130 1135 1140			3611
gga ttt gag tta ctg ctg gat tac tat gct ggc aag aac gcc agc Gly Phe Glu Leu Leu Leu Asp Tyr Tyr Ala Gly Lys Asn Ala Ser 1145 1150 1155			3656
agc att ctg aac tca gcc atg caa cag gcg tgc cgg aaa agt gac Ser Ile Leu Asn Ser Ala Met Gln Gln Ala Cys Arg Lys Ser Asp 1160 1165 1170			3701
cac ctc agt gtg agg cct agc tgt ccc tct aag cag tcc agc aca His Leu Ser Val Arg Pro Ser Cys Pro Ser Lys Gln Ser Ser Thr 1175 1180 1185			3746
gag agc atc act gag gag ttc tac agg tac atg ctg agg gac atc Glu Ser Ile Thr Glu Glu Phe Tyr Arg Tyr Met Leu Arg Asp Ile 1190 1195 1200			3791
gaa aga gac agc aga gaa agt gcc tcc tcc aga cgg agc agc cag Glu Arg Asp Ser Arg Glu Ser Ala Ser Ser Arg Arg Ser Ser Gln 1205 1210 1215			3836
gat tgg aca gcc ggc ctg ctg tct cct tct ctg cga tcc cca gtg Asp Trp Thr Ala Gly Leu Leu Ser Pro Ser Leu Arg Ser Pro Val 1220 1225 1230			3881
tgc cac aga cag tcc tcc atg cca gac agc aga tcc cca tgc tcc Cys His Arg Gln Ser Ser Met Pro Asp Ser Arg Ser Pro Cys Ser 1235 1240 1245			3926
agg ctg aca gtg aat gtg ccc atc aaa gcc aac tct tta gat ggc Arg Leu Thr Val Asn Val Pro Ile Lys Ala Asn Ser Leu Asp Gly 1250 1255 1260			3971
ttt gct cag aac tgc cca caa gat ttc cta agc gtg cag ccg gtc Phe Ala Gln Asn Cys Pro Gln Asp Phe Leu Ser Val Gln Pro Val 1265 1270 1275			4016
agt agc gcg tcc tca tcc ggt ctc tgc aaa tct gac tct tgc ttg Ser Ser Ala Ser Ser Ser Gly Leu Cys Lys Ser Asp Ser Cys Leu 1280 1285 1290			4061
tat cgg aga ggt ggg act gac cac atc acc aac atg tta att cat Tyr Arg Arg Gly Gly Thr Asp His Ile Thr Asn Met Leu Ile His 1295 1300 1305			4106
gaa acg tgg gct agc tcc att gag gct ctc atg cgc aag aac aaa Glu Thr Trp Ala Ser Ser Ile Glu Ala Leu Met Arg Lys Asn Lys 1310 1315 1320			4151
atc att gtg gat gat gca gag gaa gct gac act gag cct gtt tct			4196

16U 100 PCT.ST25													
Ile Ile Val	Asp Asp Ala Glu	Glu	Ala Asp Thr Glu	Pro	Val Ser								
1325	1330	1335											
ggt ggc tct	ccc tcg caa gca	gag	aag tgt gca aat	aga	tta gct	4241							
Gly Gly Ser	Pro Ser Gln Ala	Glu	Lys Cys Ala Asn	Arg	Leu Ala								
1340	1345	1350											
gcg agc agg	atg tgc agt ggg	cca	act ctg ctt gtt	cag	gag tct	4286							
Ala Ser Arg	Met Cys Ser Gly	Pro	Thr Leu Leu Val	Gln	Glu Ser								
1355	1360	1365											
ctc gat tgc	ccg agg aaa gac	tct	gtt acc gaa tgt	aaa	cag ccc	4331							
Leu Asp Cys	Pro Arg Lys Asp	Ser	Val Thr Glu Cys	Lys	Gln Pro								
1370	1375	1380											
cca gtg tca	tct ttg agc aaa	act	gct tct ctt aca	aac	cac agc	4376							
Pro Val Ser	Ser Leu Ser Lys	Thr	Ala Ser Leu Thr	Asn	His Ser								
1385	1390	1395											
cct tta gat	tct aaa aaa gaa	act	tcc tcg tgc cag	gac	cct gta	4421							
Pro Leu Asp	Ser Lys Lys Glu	Thr	Ser Ser Cys Gln	Asp	Pro Val								
1400	1405	1410											
cca ata aac	cac aaa agg cga	tca	ctt tgc tcg agg	gaa	gtg cct	4466							
Pro Ile Asn	His Lys Arg Arg	Ser	Leu Cys Ser Arg	Glu	Val Pro								
1415	1420	1425											
ttg att cag	att gaa aca gat	cag	aga gaa gcc tgt	gct	ggg gaa	4511							
Leu Ile Gln	Ile Glu Thr Asp	Gln	Arg Glu Ala Cys	Ala	Gly Glu								
1430	1435	1440											
cct gaa ccc	ttc ctt tcc aaa	agc	agc ctc cta gag	gaa	gca gaa	4556							
Pro Glu Pro	Phe Leu Ser Lys	Ser	Ser Leu Leu Glu	Glu	Ala Glu								
1445	1450	1455											
ggg cat tcg	aat gac aaa aac	atc	cca gat gtg gtg	aga	ggt gga	4601							
Gly His Ser	Asn Asp Lys Asn	Ile	Pro Asp Val Val	Arg	Gly Gly								
1460	1465	1470											
gac aca gcc	gtg agc gct tgt	caa	atc cat agt gac	agc	ctt gat	4646							
Asp Thr Ala	Val Ser Ala Cys	Gln	Ile His Ser Asp	Ser	Leu Asp								
1475	1480	1485											
acc aga gat	gta cca gag gct	gaa	gcc tcc aca gaa	gcc	aga gcc	4691							
Thr Arg Asp	Val Pro Glu Ala	Glu	Ala Ser Thr Glu	Ala	Arg Ala								
1490	1495	1500											
ccc gat gag	gcc ccc aac cct	cca	agc agc agc gag	gag	agc aca	4736							
Pro Asp Glu	Ala Pro Asn Pro	Pro	Ser Ser Ser Glu	Glu	Ser Thr								
1505	1510	1515											
ggc agc tgg	acc cag ctt gcc	aat	gag gaa gac aac	cca	gat gac	4781							
Gly Ser Trp	Thr Gln Leu Ala	Asn	Glu Glu Asp Asn	Pro	Asp Asp								
1520	1525	1530											
aca agt agc	ttt ctc cag ctc	agt	gag cga tcc atg	agc	aat ggc	4826							
Thr Ser Ser	Phe Leu Gln Leu	Ser	Glu Arg Ser Met	Ser	Asn Gly								
1535	1540	1545											
aac agt agt	gcc act agc agt	ctt	ggc att atg gat	ctg	gac att	4871							
Asn Ser Ser	Ala Thr Ser Ser	Leu	Gly Ile Met Asp	Leu	Asp Ile								
1550	1555	1560											
tat cag gaa	agc atg cca tct	tct	ccc atg att aat	gaa	tta gta	4916							
Tyr Gln Glu	Ser Met Pro Ser	Ser	Pro Met Ile Asn	Glu	Leu Val								
1565	1570	1575											
gaa gaa aag	aag att ctt aaa	gga	cag tca gaa agc	aca	gag gca	4961							
Glu Glu Lys	Lys Ile Leu Lys	Gly	Gln Ser Glu Ser	Thr	Glu Ala								
1580	1585	1590											
cct gca tct	gga ccg cct acg	gga	aca gcc agc ccc	cag	agg agc	5006							
Pro Ala Ser	Gly Pro Pro Thr	Gly	Thr Ala Ser Pro	Gln	Arg Ser								
1595	1600	1605											
ctg ctg gtg	atc aac ttt gac	ctg	gag cca gag tgt	cca	gat gcc	5051							
Leu Leu Val	Ile Asn Phe Asp	Leu	Glu Pro Glu Cys	Pro	Asp Ala								
1610	1615	1620											

160 100 PCT.ST25

gag ctc cga	gcc act ctg cag tgg	ata gct gcc tct gaa	ctg ggg	5096		
Glu Leu Arg	Ala Thr Leu Gln Trp	Ile Ala Ala Ser Glu	Leu Gly			
1625	1630	1635				
att ccc acc	atc tac ttt aag aaa	tct cag gaa aac aga	att gaa	5141		
Ile Pro Thr	Ile Tyr Phe Lys Lys	Ser Gln Glu Asn Arg	Ile Glu			
1640	1645	1650				
aag ttt cta	gat gtc gtg cag ctg	gtt cat cgg aag tcc	tgg aaa	5186		
Lys Phe Leu	Asp Val Val Gln Leu	Val His Arg Lys Ser	Trp Lys			
1655	1660	1665				
gtg ggt gat	atc ttc cat gca gtt	gtc cag tac tgc aaa	atg cat	5231		
Val Gly Asp	Ile Phe His Ala Val	Val Gln Tyr Cys Lys	Met His			
1670	1675	1680				
gag gag cag	aag gat ggg aga ctg	agt ctc ttt gac tgg	ctc ttg	5276		
Glu Glu Gln	Lys Asp Gly Arg Leu	Ser Leu Phe Asp Trp	Leu Leu			
1685	1690	1695				
gaa ctg gga	taa taaggcagtc	tgccgtatag atcattcctt	ccctttattc	5328		
Glu Leu Gly						
1700						
caacttagat	tacagtgggt	tggtctaaat	gctctaaaca	ttctcaaac	atcacatcac	5388
attagcagaa	ctataaaaa	aaatctgcta	ctcagatcca	ctgcatacag	aataagtcag	5448
aggaaaagca	aaatataggt	ctgtccaaat	tcatacaact	tgtgggtgag	ttccaaagag	5508
cttggttag	aagggtgga	caaagagaga	attcaatggg	gcccaaatta	gaatgcttat	5568
aatgagaccc	aatctccag	aaaacaacac	tcacataagt	ttaatcatat	aaaatgattt	5628
gtaatgtctc	taattagatg	aatcaactag	aaacaaactc	agtggtcaaa	ataattttta	5688
agagtattcc	gtaacctata	ttttactttt	ctgattatat	taaggggctg	ccagcccgga	5748
gaaatactta	agatatgggt	gagaaatccc	cagactttta	tacaaaagat	ttccactttc	5808
aaatcaatgt	cagtagacat	tgataaaagt	atagcagcat	cctctactga	ggtgatttca	5868
tttattccct	gcagcccact	gataaatatc	tcacttctcc	caaatagtat	gtggactccc	5928
agctaagcag	aaaactattg	tcattcaact	gaagaggaag	ataaaagatt	gtcttggttc	5988
catcactgta	ttacttgtgt	aacatgatta	cataattctt	atcctaagag	aaagctttca	6048
tatttaaaaa	aaagtctttt	cagataaaat	ctgcttgtgt	cttgaataat	atgaaataca	6108
aacttttact	ttattttatt	gtaaattata	aagagattat	tgtcttaaat	aatatattga	6168
gttagcttca	agcttccata	aatatgaaga	gattgtgtgc	taaagtcaca	tattgacatt	6228
gagctcagtg	gcctgtttca	tcacgtatgt	gctgctacct	gtacagcaga	catgccgctc	6288
cagtgcatt	tataatgaca	gaagcagggt	aatgggtcttg	tgtttgacat	gatcagttag	6348
gatcatagac	tttcctgac	tcgtagatat	tagccttgaa	ttgggggaaa	agaagacttt	6408
gacacatttt	agttatttta	ataacagaga	tttactcttt	tgaaaaataa	aggtatctaa	6468
tgtctcccta	ataagtcttc	tttccttcca	actaaatgac	ctacacggac	ttttattttc	6528
ttgatcaaa	agggttttat	taaggacttc	tgataacta	tacttttact	ctatttttaa	6588
agatcacaaa	gtaattttta	atgtgaacag	gttcccatat	catgaatgct	ggcctcacct	6648
tctctatcat	ccacattttg	aaatgcaaag	aaagctccct	tgtaagccat	acttccctcc	6708
ccactcccat	cctaggatag	ttgcccagtg	ctcattaggt	atttcttatt	cagatagttc	6768
aaatttaggt	tattatgctt	aatttgacac	attaactaaa	tgcccagttt	taaaatatat	6828
ccatcaattc	acgctgaaat	gtgcttcttt	gtgctatcaa	atggaataga	atacacttat	6888
tttttaaa	atccagaat	actgtgtgta	gacttttggt	gtgctcaaat	aaatgtttac	6948
ttatcttaca	aagctcaaat	actggattgt	aaccatgtga	tgaagttatc	tatgtttgac	7008

160 100 PCT.ST25

ctaacattgc aaattaatca ataaatctct gttgtcaaaa aaaaaaaaaa aaaa

7062

<210> 39
 <211> 1700
 <212> PRT
 <213> Homo sapiens

<400> 39

Met Asp Gly Asn Ser Leu Leu Ser Val Pro Ser Asn Leu Glu Ser Ser
 1 5 10 15

Arg Met Tyr Asp Val Leu Glu Pro Gln Gln Gly Arg Gly Cys Gly Ser
 20 25 30

Ser Gly Ser Gly Pro Gly Asn Ser Ile Thr Ala Cys Lys Lys Val Leu
 35 40 45

Arg Ser Asn Ser Leu Leu Glu Ser Thr Asp Tyr Trp Leu Gln Asn Gln
 50 55 60

Arg Met Pro Cys Gln Ile Gly Phe Val Glu Asp Lys Ser Glu Asn Cys
 65 70 75 80

Ala Ser Val Cys Phe Val Asn Leu Asp Val Asn Lys Asp Glu Cys Ser
 85 90 95

Thr Glu His Leu Gln Gln Lys Leu Val Asn Val Ser Pro Asp Leu Pro
 100 105 110

Lys Leu Ile Ser Ser Met Asn Val Gln Gln Pro Lys Glu Asn Glu Ile
 115 120 125

Val Val Leu Ser Gly Leu Ala Ser Gly Asn Leu Gln Ala Asp Phe Glu
 130 135 140

Val Ser Gln Cys Pro Trp Leu Pro Asp Ile Cys Leu Val Gln Cys Ala
 145 150 155 160

Arg Gly Asn Arg Pro Asn Ser Thr Asn Cys Ile Ile Phe Glu Ile Asn
 165 170 175

Lys Phe Leu Ile Gly Leu Glu Leu Val Gln Glu Arg Gln Leu His Leu
 180 185 190

Glu Thr Asn Ile Leu Lys Leu Glu Asp Asp Thr Asn Cys Ser Leu Ser
 195 200 205

Ser Ile Glu Glu Asp Phe Leu Thr Ala Ser Glu His Leu Glu Glu Glu
 210 215 220

Ser Glu Val Asp Glu Ser Arg Asn Asp Tyr Glu Asn Ile Asn Val Ser
 225 230 235 240

Ala Asn Val Leu Glu Ser Lys Gln Leu Lys Gly Ala Thr Gln Val Glu
 245 250 255

Trp Asn Cys Asn Lys Glu Lys Trp Leu Tyr Ala Leu Glu Asp Lys Tyr
 260 265 270

160 100 PCT.ST25

Ile Asn Lys Tyr Pro Thr Pro Leu Ile Lys Thr Glu Arg Ser Pro Glu
275 280 285

Asn Leu Thr Lys Asn Thr Ala Leu Gln Ser Leu Asp Pro Ser Ala Lys
290 295 300

Pro Ser Gln Trp Lys Arg Glu Ala Val Gly Asn Gly Arg Gln Ala Thr
305 310 315 320

His Tyr Tyr His Ser Glu Ala Phe Lys Gly Gln Met Glu Lys Ser Gln
325 330 335

Ala Leu Tyr Ile Pro Lys Asp Ala Tyr Phe Ser Met Met Asp Lys Asp
340 345 350

Val Pro Ser Ala Cys Ala Val Ala Glu Gln Arg Ser Asn Leu Asn Pro
355 360 365

Gly Asp His Glu Asp Thr Arg Asn Ala Leu Pro Pro Arg Gln Asp Gly
370 375 380

Glu Val Thr Thr Gly Lys Tyr Ala Thr Asn Leu Ala Glu Ser Val Leu
385 390 395 400

Gln Asp Ala Phe Ile Arg Leu Ser Gln Ser Gln Ser Thr Leu Pro Gln
405 410 415

Glu Ser Ala Val Ser Val Ser Val Gly Ser Ser Leu Leu Pro Ser Cys
420 425 430

Tyr Ser Thr Lys Asp Thr Val Val Ser Arg Ser Trp Asn Glu Leu Pro
435 440 445

Lys Ile Val Val Val Gln Ser Pro Asp Gly Ser Asp Ala Ala Pro Gln
450 455 460

Pro Gly Ile Ser Ser Trp Pro Glu Met Glu Val Ser Val Glu Thr Ser
465 470 475 480

Ser Ile Leu Ser Gly Glu Asn Ser Ser Arg Gln Pro Gln Ser Ala Leu
485 490 495

Glu Val Ala Leu Ala Cys Ala Ala Thr Val Ile Gly Thr Ile Ser Ser
500 505 510

Pro Gln Ala Thr Glu Arg Leu Lys Met Glu Gln Val Val Ser Asn Phe
515 520 525

Pro Pro Gly Ser Ser Gly Ala Leu Gln Thr Gln Ala Pro Gln Gly Leu
530 535 540

Lys Glu Pro Ser Ile Asn Glu Tyr Ser Phe Pro Ser Ala Leu Cys Gly
545 550 555 560

Met Thr Gln Val Ala Ser Ala Val Ala Val Cys Gly Leu Gly Glu Arg
565 570 575

Glu Glu Val Thr Cys Ser Val Ala Pro Ser Gly Ser Leu Pro Pro Ala
580 585 590

160 100 PCT.ST25

Ala Glu Ala Ser Glu Ala Met Pro Pro Leu Cys Gly Leu Ala Ser Met
 595 600 605
 Glu Leu Gly Lys Glu Ala Ile Ala Glu Gly Leu Leu Lys Glu Ala Ala
 610 615 620
 Leu Val Leu Thr Arg Pro Asn Thr Tyr Ser Ser Ile Gly Asp Phe Leu
 625 630 635 640
 Asp Ser Met Asn Arg Arg Ile Met Glu Thr Ala Ser Lys Ser Gln Thr
 645 650 655
 Leu Cys Ser Glu Asn Val Val Arg Asn Glu Leu Ala His Thr Leu Ser
 660 665 670
 Asn Val Ile Leu Arg His Ser Ile Asp Glu Val His His Lys Asn Met
 675 680 685
 Ile Ile Asp Pro Asn Asp Asn Arg His Ser Ser Glu Ile Leu Asp Thr
 690 695 700
 Leu Met Glu Ser Thr Asn Gln Leu Leu Leu Asp Val Ile Cys Phe Thr
 705 710 715 720
 Phe Lys Lys Met Ser His Ile Val Arg Leu Gly Glu Cys Pro Ala Val
 725 730 735
 Leu Ser Lys Glu Thr Ile Arg Arg Arg Glu Thr Glu Pro Ser Cys Gln
 740 745 750
 Pro Ser Asp Pro Gly Ala Ser Gln Ala Trp Thr Lys Ala Thr Glu Ser
 755 760 765
 Ser Ser Ser Ser Pro Leu Ser Asn Ser His Asn Thr Ser Leu Val Ile
 770 775 780
 Asn Asn Leu Val Asp Gly Met Tyr Ser Lys Gln Asp Lys Gly Gly Val
 785 790 795 800
 Arg Pro Gly Leu Phe Lys Asn Pro Thr Leu Gln Ser Gln Leu Ser Arg
 805 810 815
 Ser His Arg Val Pro Asp Ser Ser Thr Ala Thr Thr Ser Ser Lys Glu
 820 825 830
 Ile Tyr Leu Lys Gly Ile Ala Gly Glu Asp Thr Lys Ser Pro His His
 835 840 845
 Ser Glu Asn Glu Cys Arg Ala Ser Ser Glu Gly Gln Arg Ser Pro Thr
 850 855 860
 Val Ser Arg Ser Arg Ser Gly Ser Gln Glu Ala Glu Glu Ser Ile His
 865 870 875 880
 Pro Asn Thr Gln Glu Lys Tyr Asn Cys Ala Thr Ser Arg Ile Asn Glu
 885 890 895
 Val Gln Val Asn Leu Ser Leu Leu Gly Asp Asp Leu Leu Leu Pro Ala
 900 905 910

16U 100 PCT.ST25

Gln Ser Thr Leu Gln Thr Lys His Pro Asp Ile Tyr Cys Ile Thr Asp
 915 920 925

Phe Ala Glu Glu Leu Ala Asp Thr Val Val Ser Met Ala Thr Glu Ile
 930 935 940

Ala Ala Ile Cys Leu Asp Asn Ser Ser Gly Lys Gln Pro Trp Phe Cys
 945 950 955 960

Ala Trp Lys Arg Gly Ser Glu Phe Leu Met Thr Pro Asn Val Pro Cys
 965 970 975

Arg Ser Leu Lys Arg Lys Lys Glu Ser Gln Gly Ser Gly Thr Ala Val
 980 985 990

Arg Lys His Lys Pro Pro Arg Leu Ser Glu Ile Lys Arg Lys Thr Asp
 995 1000 1005

Glu His Pro Glu Leu Lys Glu Lys Leu Met Asn Arg Val Val Asp
 1010 1015 1020

Glu Ser Met Asn Leu Glu Asp Val Pro Asp Ser Val Asn Leu Phe
 1025 1030 1035

Ala Asn Glu Val Ala Ala Lys Ile Met Asn Leu Thr Glu Phe Ser
 1040 1045 1050

Met Val Asp Gly Met Trp Gln Ala Gln Gly Tyr Pro Arg Asn Arg
 1055 1060 1065

Leu Leu Ser Gly Asp Arg Trp Ser Arg Leu Lys Ala Ser Ser Cys
 1070 1075 1080

Glu Ser Ile Pro Glu Glu Asp Ser Glu Ala Arg Ala Tyr Val Asn
 1085 1090 1095

Ser Leu Gly Leu Met Ser Thr Leu Ser Gln Pro Val Ser Arg Ala
 1100 1105 1110

Ser Ser Val Ser Lys Gln Ser Ser Cys Glu Ser Ile Thr Asp Glu
 1115 1120 1125

Phe Ser Arg Phe Met Val Asn Gln Met Glu Asn Glu Gly Arg Gly
 1130 1135 1140

Phe Glu Leu Leu Leu Asp Tyr Tyr Ala Gly Lys Asn Ala Ser Ser
 1145 1150 1155

Ile Leu Asn Ser Ala Met Gln Gln Ala Cys Arg Lys Ser Asp His
 1160 1165 1170

Leu Ser Val Arg Pro Ser Cys Pro Ser Lys Gln Ser Ser Thr Glu
 1175 1180 1185

Ser Ile Thr Glu Glu Phe Tyr Arg Tyr Met Leu Arg Asp Ile Glu
 1190 1195 1200

Arg Asp Ser Arg Glu Ser Ala Ser Ser Arg Arg Ser Ser Gln Asp

1205	1210	160 100 PCT.ST25 1215
Trp Thr Ala Gly Leu Leu Ser 1220 1225	Pro Ser Leu Arg Ser	Pro Val Cys 1230
His Arg Gln Ser Ser Met Pro 1235 1240	Asp Ser Arg Ser	Pro Cys Ser Arg 1245
Leu Thr Val Asn Val Pro Ile 1250 1255	Lys Ala Asn Ser	Leu Asp Gly Phe 1260
Ala Gln Asn Cys Pro Gln Asp 1265 1270	Phe Leu Ser Val	Gln Pro Val Ser 1275
Ser Ala Ser Ser Ser Gly Leu 1280 1285	Cys Lys Ser Asp	Ser Cys Leu Tyr 1290
Arg Arg Gly Gly Thr Asp His 1295 1300	Ile Thr Asn Met	Leu Ile His Glu 1305
Thr Trp Ala Ser Ser Ile Glu 1310 1315	Ala Leu Met Arg	Lys Asn Lys Ile 1320
Ile Val Asp Asp Ala Glu Glu 1325 1330	Ala Asp Thr Glu	Pro Val Ser Gly 1335
Gly Ser Pro Ser Gln Ala Glu 1340 1345	Lys Cys Ala Asn	Arg Leu Ala Ala 1350
Ser Arg Met Cys Ser Gly Pro 1355 1360	Thr Leu Leu Val	Gln Glu Ser Leu 1365
Asp Cys Pro Arg Lys Asp Ser 1370 1375	Val Thr Glu Cys	Lys Gln Pro Pro 1380
Val Ser Ser Leu Ser Lys Thr 1385 1390	Ala Ser Leu Thr	Asn His Ser Pro 1395
Leu Asp Ser Lys Lys Glu Thr 1400 1405	Ser Ser Cys Gln	Asp Pro Val Pro 1410
Ile Asn His Lys Arg Arg Ser 1415 1420	Leu Cys Ser Arg	Glu Val Pro Leu 1425
Ile Gln Ile Glu Thr Asp Gln 1430 1435	Arg Glu Ala Cys	Ala Gly Glu Pro 1440
Glu Pro Phe Leu Ser Lys Ser 1445 1450	Ser Leu Leu Glu	Glu Ala Glu Gly 1455
His Ser Asn Asp Lys Asn Ile 1460 1465	Pro Asp Val Val	Arg Gly Gly Asp 1470
Thr Ala Val Ser Ala Cys Gln 1475 1480	Ile His Ser Asp	Ser Leu Asp Thr 1485
Arg Asp Val Pro Glu Ala Glu 1490 1495	Ala Ser Thr Glu	Ala Arg Ala Pro 1500

16U 100 PCT.ST25
 Asp Glu Ala Pro Asn Pro Pro Ser Ser Ser Glu Glu Ser Thr Gly
 1505 1510 1515

Ser Trp Thr Gln Leu Ala Asn Glu Glu Asp Asn Pro Asp Asp Thr
 1520 1525 1530

Ser Ser Phe Leu Gln Leu Ser Glu Arg Ser Met Ser Asn Gly Asn
 1535 1540 1545

Ser Ser Ala Thr Ser Ser Leu Gly Ile Met Asp Leu Asp Ile Tyr
 1550 1555 1560

Gln Glu Ser Met Pro Ser Ser Pro Met Ile Asn Glu Leu Val Glu
 1565 1570 1575

Glu Lys Lys Ile Leu Lys Gly Gln Ser Glu Ser Thr Glu Ala Pro
 1580 1585 1590

Ala Ser Gly Pro Pro Thr Gly Thr Ala Ser Pro Gln Arg Ser Leu
 1595 1600 1605

Leu Val Ile Asn Phe Asp Leu Glu Pro Glu Cys Pro Asp Ala Glu
 1610 1615 1620

Leu Arg Ala Thr Leu Gln Trp Ile Ala Ala Ser Glu Leu Gly Ile
 1625 1630 1635

Pro Thr Ile Tyr Phe Lys Lys Ser Gln Glu Asn Arg Ile Glu Lys
 1640 1645 1650

Phe Leu Asp Val Val Gln Leu Val His Arg Lys Ser Trp Lys Val
 1655 1660 1665

Gly Asp Ile Phe His Ala Val Val Gln Tyr Cys Lys Met His Glu
 1670 1675 1680

Glu Gln Lys Asp Gly Arg Leu Ser Leu Phe Asp Trp Leu Leu Glu
 1685 1690 1695

Leu Gly
 1700

<210> 40
 <211> 25
 <212> DNA
 <213> Homo sapiens

<400> 40
 cgtcaggaat gaactggcac atacc

25

<210> 41
 <211> 24
 <212> DNA
 <213> Homo sapiens; Homo sapiens

<400> 41
 ctaagtggag agctgctgga ggat

24

<210> 42
 <211> 1302
 <212> PRT
 <213> Homo sapiens

16U 100 PCT.ST25

<400> 42

Asp His Glu Asp Thr Arg Asn Ala Leu Pro Pro Arg Gln Asp Gly Glu
 1 5 10 15
 Val Thr Thr Gly Lys Tyr Ala Thr Asn Leu Ala Glu Ser Val Leu Gln
 20 25 30
 Asp Ala Phe Ile Arg Leu Ser Gln Ser Gln Ser Thr Leu Pro Gln Glu
 35 40 45
 Ser Ala Val Ser Val Ser Val Gly Ser Ser Leu Leu Pro Ser Cys Tyr
 50 55 60
 Ser Thr Lys Asp Thr Val Val Ser Arg Ser Trp Asn Glu Leu Pro Lys
 65 70 75 80
 Ile Val Val Val Gln Ser Pro Asp Gly Ser Asp Ala Ala Pro Gln Pro
 85 90 95
 Gly Ile Ser Ser Trp Pro Glu Met Glu Val Ser Val Glu Thr Ser Ser
 100 105 110
 Ile Leu Ser Gly Glu Asn Ser Ser Arg Gln Pro Gln Ser Ala Leu Glu
 115 120 125
 Val Ala Leu Ala Cys Ala Ala Thr Val Ile Gly Thr Ile Ser Ser Pro
 130 135 140
 Gln Ala Thr Glu Arg Leu Lys Met Glu Gln Val Val Ser Asn Phe Pro
 145 150 155 160
 Pro Gly Ser Ser Gly Ala Leu Gln Thr Gln Ala Pro Gln Gly Leu Lys
 165 170 175
 Glu Pro Ser Ile Asn Glu Tyr Ser Phe Pro Ser Ala Leu Cys Gly Met
 180 185 190
 Thr Gln Val Ala Ser Ala Val Ala Val Cys Gly Leu Gly Glu Arg Glu
 195 200 205
 Glu Val Thr Cys Ser Val Ala Pro Ser Gly Ser Leu Pro Pro Ala Ala
 210 215 220
 Glu Ala Ser Glu Ala Met Pro Pro Leu Cys Gly Leu Ala Ser Met Glu
 225 230 235 240
 Leu Gly Lys Glu Ala Ile Ala Glu Gly Leu Leu Lys Glu Ala Ala Leu
 245 250 255
 Val Leu Thr Arg Pro Asn Thr Tyr Ser Ser Ile Gly Asp Phe Leu Asp
 260 265 270
 Ser Met Asn Arg Arg Ile Met Glu Thr Ala Ser Lys Ser Gln Thr Leu
 275 280 285
 Cys Ser Glu Asn Val Val Arg Asn Glu Leu Ala His Thr Leu Ser Asn
 290 295 300
 Val Ile Leu Arg His Ser Ile Asp Glu Val His His Lys Asn Met Ile

Page 55

16U 100 PCT.ST25

Lys His Lys Pro Pro Arg Leu Ser Glu Ile Lys Arg Lys Thr Asp Glu
625 630 635 640

His Pro Glu Leu Lys Glu Lys Leu Met Asn Arg Val Val Asp Glu Ser
645 650 655

Met Asn Leu Glu Asp Val Pro Asp Ser Val Asn Leu Phe Ala Asn Glu
660 665 670

Val Ala Ala Lys Ile Met Asn Leu Thr Glu Phe Ser Met Val Asp Gly
675 680 685

Met Trp Gln Ala Gln Gly Tyr Pro Arg Asn Arg Leu Leu Ser Gly Asp
690 695 700

Arg Trp Ser Arg Leu Lys Ala Ser Ser Cys Glu Ser Ile Pro Glu Glu
705 710 715 720

Asp Ser Glu Ala Arg Ala Tyr Val Asn Ser Leu Gly Leu Met Ser Thr
725 730 735

Leu Ser Gln Pro Val Ser Arg Ala Ser Ser Val Ser Lys Gln Ser Ser
740 745 750

Cys Glu Ser Ile Thr Asp Glu Phe Ser Arg Phe Met Val Asn Gln Met
755 760 765

Glu Asn Glu Gly Arg Gly Phe Glu Leu Leu Leu Asp Tyr Tyr Ala Gly
770 775 780

Lys Asn Ala Ser Ser Ile Leu Asn Ser Ala Met Gln Gln Ala Cys Arg
785 790 795 800

Lys Ser Asp His Leu Ser Val Arg Pro Ser Cys Pro Ser Lys Gln Ser
805 810 815

Ser Thr Glu Ser Ile Thr Glu Glu Phe Tyr Arg Tyr Met Leu Arg Asp
820 825 830

Ile Glu Arg Asp Ser Arg Glu Ser Ala Ser Ser Arg Arg Ser Ser Gln
835 840 845

Asp Trp Thr Ala Gly Leu Leu Ser Pro Ser Leu Arg Ser Pro Val Cys
850 855 860

His Arg Gln Ser Ser Met Pro Asp Ser Arg Ser Pro Cys Ser Arg Leu
865 870 875 880

Thr Val Asn Val Pro Ile Lys Ala Asn Ser Leu Asp Gly Phe Ala Gln
885 890 895

Asn Cys Pro Gln Asp Phe Leu Ser Val Gln Pro Val Ser Ser Ala Ser
900 905 910

Ser Ser Gly Leu Cys Lys Ser Asp Ser Cys Leu Tyr Arg Arg Gly Gly
915 920 925

Thr Asp His Ile Thr Asn Met Leu Ile His Glu Thr Trp Ala Ser Ser
930 935 940

16U 100 PCT.ST25

Ile Glu Ala Leu Met Arg Lys Asn Lys Ile Ile Val Asp Asp Ala Glu
 945 950 955 960
 Glu Ala Asp Thr Glu Pro Val Ser Gly Gly Ser Pro Ser Gln Ala Glu
 965 970 975
 Lys Cys Ala Asn Arg Leu Ala Ala Ser Arg Met Cys Ser Gly Pro Thr
 980 985 990
 Leu Leu Val Gln Glu Ser Leu Asp Cys Pro Arg Lys Asp Ser Val Thr
 995 1000 1005
 Glu Cys Lys Gln Pro Pro Val Ser Ser Leu Ser Lys Thr Ala Ser
 1010 1015 1020
 Leu Thr Asn His Ser Pro Leu Asp Ser Lys Lys Glu Thr Ser Ser
 1025 1030 1035
 Cys Gln Asp Pro Val Pro Ile Asn His Lys Arg Arg Ser Leu Cys
 1040 1045 1050
 Ser Arg Glu Val Pro Leu Ile Gln Ile Glu Thr Asp Gln Arg Glu
 1055 1060 1065
 Ala Cys Ala Gly Glu Pro Glu Pro Phe Leu Ser Lys Ser Ser Leu
 1070 1075 1080
 Leu Glu Glu Ala Glu Gly His Ser Asn Asp Lys Asn Ile Pro Asp
 1085 1090 1095
 Val Val Arg Gly Gly Asp Thr Ala Val Ser Ala Cys Gln Ile His
 1100 1105 1110
 Ser Asp Ser Leu Asp Thr Arg Asp Val Pro Glu Ala Glu Ala Ser
 1115 1120 1125
 Thr Glu Ala Arg Ala Pro Asp Glu Ala Pro Asn Pro Pro Ser Ser
 1130 1135 1140
 Ser Glu Glu Ser Thr Gly Ser Trp Thr Gln Leu Ala Asn Glu Glu
 1145 1150 1155
 Asp Asn Pro Asp Asp Thr Ser Ser Phe Leu Gln Leu Ser Glu Arg
 1160 1165 1170
 Ser Met Ser Glu Leu Val Glu Glu Lys Lys Ile Leu Lys Gly Gln
 1175 1180 1185
 Ser Glu Ser Thr Glu Ala Pro Ala Ser Gly Pro Pro Thr Gly Thr
 1190 1195 1200
 Ala Ser Pro Gln Arg Ser Leu Leu Val Ile Asn Phe Asp Leu Glu
 1205 1210 1215
 Pro Glu Cys Pro Asp Ala Glu Leu Arg Ala Thr Leu Gln Trp Ile
 1220 1225 1230
 Ala Ala Ser Glu Leu Gly Ile Pro Thr Ile Tyr Phe Lys Lys Ser
 1235 1240 1245

16U 100 PCT.ST25

Gln Glu Asn Arg Ile Glu Lys Phe Leu Asp Val Val Gln Leu Val
1250 1255 1260

His Arg Lys Ser Trp Lys Val Gly Asp Ile Phe His Ala Val Val
1265 1270 1275

Gln Tyr Cys Lys Met His Glu Glu Gln Lys Asp Gly Arg Leu Ser
1280 1285 1290

Leu Phe Asp Trp Leu Leu Glu Leu Gly
1295 1300

<210> 43
<211> 3369
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (250)..(2793)
<223>

<400> 43
gctggatcaa gctgtgaacg tgatttgctg gaagctggtt gacgatgtgt cacactgtgt 60
aagggaatcg catggagatg ggcattccga actgttaatg gggacatggg actccagttg 120
tctctgatca cttgtgtgga ttttcctggc gtagaacgac agaagccgct agtaagtcgc 180
caagacctac agcaggaatt ctgcaccaaa gggcataaaa tcttggtatt ttaatttgca 240
tctgggaga atg tct gag caa gga gac ctg aat cag gca ata gca gag gaa 291
Met Ser Glu Gln Gly Asp Leu Asn Gln Ala Ile Ala Glu Glu
1 5 10
gga ggg act gag cag gag acg gcc act cca gag aac ggc att gtt aaa 339
Gly Gly Thr Glu Gln Glu Thr Ala Thr Pro Glu Asn Gly Ile Val Lys
15 20 25 30
tca gaa agt ctg gat gaa gag gag aaa ctg gaa ctg cag agg cgg ctg 387
Ser Glu Ser Leu Asp Glu Glu Glu Lys Leu Glu Leu Gln Arg Arg Leu
35 40 45
gag gct cag aat caa gaa aga aga aaa tcc aag tca gga gca gga aaa 435
Glu Ala Gln Asn Gln Glu Arg Arg Lys Ser Lys Ser Gly Ala Gly Lys
50 55 60
ggt aaa ctg act cgc agt ctt gct gtc tgt gag gaa tct tct gcc aga 483
Gly Lys Leu Thr Arg Ser Leu Ala Val Cys Glu Glu Ser Ser Ala Arg
65 70 75
cca gga ggt gaa agt ctt cag gat cag gaa tca att cat tta cag ctt 531
Pro Gly Gly Glu Ser Leu Gln Asp Gln Glu Ser Ile His Leu Gln Leu
80 85 90
tcc agt ttt tcc agc ctg caa gag gag gat aaa tct agg aaa gat gac 579
Ser Ser Phe Ser Ser Leu Gln Glu Glu Asp Lys Ser Arg Lys Asp Asp
95 100 105 110
tct gaa aga gaa aaa gaa aag gat aaa aac aaa gat aaa acc tct gaa 627
Ser Glu Arg Glu Lys Glu Lys Asp Lys Asn Lys Asp Lys Thr Ser Glu
115 120 125
aaa ccc aag atc aga atg tta tca aaa gat tgc agc caa gaa tac acg 675
Lys Pro Lys Ile Arg Met Leu Ser Lys Asp Cys Ser Gln Glu Tyr Thr
130 135 140
gat tct aca ggc ata gac tta cac gag ttt ctg att aac aca tta aag 723
Asp Ser Thr Gly Ile Asp Leu His Glu Phe Leu Ile Asn Thr Leu Lys
145 150 155
aat aat tcc agg gac agg atg ata ctt ttg aaa atg gag cag gaa att 771
Asn Asn Ser Arg Asp Arg Met Ile Leu Leu Lys Met Glu Gln Glu Ile

160 100 PCT.ST25																
160						165						170				
att Ile 175	gat Asp	ttc Phe	att Ile	gct Ala	gac Asp 180	aac Asn	aat Asn	aat Asn	cat His	tat Tyr 185	aaa Lys	aag Lys	ttc Phe	cct Pro	cag Gln 190	819
atg Met	tca Ser	tcg Ser	tat Tyr	cag Gln 195	agg Arg	atg Met	ctt Leu	gtc Val	cat His 200	cga Arg	gtg Val	gca Ala	gct Ala	tat Tyr 205	ttt Phe	867
gga Gly	ttg Leu	gat Asp	cac His 210	aat Asn	gtg Val	gat Asp	caa Gln	aca Thr 215	gga Gly	aaa Lys	tct Ser	gtt Val	atc Ile 220	atc Ile	aac Asn	915
aag Lys	acc Thr	agc Ser 225	agc Ser	acc Thr	aga Arg	ata Ile	cca Pro 230	gag Glu	caa Gln	agg Arg	ttt Phe	tgt Cys 235	gaa Glu	cat His	tta Leu	963
aaa Lys	gat Asp 240	gaa Glu	aaa Lys	ggt Gly	gaa Glu	gaa Glu 245	tcc Ser	cag Gln	aag Lys	cgg Arg	ttt Phe 250	atc Ile	ttg Leu	aag Lys	cga Arg	1011
gat Asp 255	aac Asn	tct Ser	agt Ser	att Ile	gat Asp 260	aaa Lys	gaa Glu	gac Asp	aat Asn	cag Gln 265	caa Gln	aac Asn	aga Arg	atg Met	cat His 270	1059
cca Pro	ttt Phe	aga Arg	gat Asp	gac Asp 275	aga Arg	cga Arg	agt Ser	aaa Lys	tca Ser 280	att Ile	gaa Glu	gag Glu	aga Arg	gaa Glu 285	gag Glu	1107
gaa Glu	tat Tyr	cag Gln	aga Arg 290	gtg Val	agg Arg	gag Glu	aga Arg	ata Ile 295	ttt Phe	gca Ala	cac His	gat Asp	tca Ser 300	gtt Val	tgc Cys	1155
tcc Ser	cag Gln	gaa Glu 305	agc Ser	ctt Leu	ttt Phe	gtg Val	gaa Glu 310	aac Asn	agt Ser	agg Arg	ctc Leu	ttg Leu 315	gaa Glu	gac Asp	agt Ser	1203
aac Asn	ata Ile 320	tgc Cys	aat Asn	gag Glu	acc Thr	tat Tyr 325	aag Lys	aaa Lys	aga Arg	cag Gln	ctc Leu 330	ttt Phe	cgg Arg	ggc Gly	aac Asn	1251
aga Arg 335	gat Asp	ggc Gly	tca Ser	ggg Gly	aga Arg 340	aca Thr	tct Ser	ggg Gly	agt Ser	cga Arg 345	cag Gln	agc Ser	agc Ser	tca Ser	gaa Glu 350	1299
aat Asn	gaa Glu	ctc Leu	aag Lys	tgg Trp 355	tct Ser	gac Asp	cac His	caa Gln	agg Arg 360	gcc Ala	tgg Trp	agc Ser	agc Ser	aca Thr 365	gac Asp	1347
tcc Ser	gac Asp	agt Ser	tcc Ser 370	aac Asn	cgc Arg	aat Asn	cta Leu	aag Lys 375	ccc Pro	gcc Ala	atg Met	acc Thr	aag Lys 380	acg Thr	gcg Ala	1395
agt Ser	ttt Phe	ggg Gly 385	ggc Gly	atc Ile	acg Thr	gtg Val	ctg Leu 390	acc Thr	agg Arg	ggt Gly	gac Asp	agc Ser 395	act Thr	tcc Ser	agt Ser	1443
act Thr 400	agg Arg	agt Ser	acc Thr	ggg Gly	aag Lys	ctg Leu 405	tcc Ser	aaa Lys	gca Ala	ggt Gly	tcc Ser 410	gag Glu	tct Ser	tcc Ser	agc Ser	1491
agt Ser 415	gca Ala	ggc Gly	tcc Ser	tca Ser	gga Gly 420	tcg Ser	ctg Leu	tcc Ser	cgc Arg	acc Thr 425	cat His	cca Pro	cct Pro	ctc Leu	cag Gln 430	1539
agc Ser	aca Thr	ccc Pro	cta Leu	gtc Val 435	tca Ser	ggt Gly	gtg Val	gca Ala	gct Ala 440	ggc Gly	tct Ser	cca Pro	ggc Gly	tgt Cys 445	gtg Val	1587
cct Pro	tat Tyr	cca Pro	gag Glu 450	aat Asn	gga Gly	ata Ile	ggg Gly	ggc Gly 455	cag Gln	gtt Val	gct Ala	ccc Pro	agc Ser 460	agc Ser	acc Thr	1635
agc Ser	tac Tyr	atc Ile 465	ctc Leu	ctt Leu	cca Pro	ctt Leu	gaa Glu 470	gct Ala	gca Ala	aca Thr	ggc Gly	atc Ile 475	ccg Pro	cct Pro	gga Gly	1683
agc	atc	ctt	ctt	aat	cca	cac	aca	gqc	cag	ccc	ttt	gtg	aat	ccc	gat	1731

16U 100 PCT.ST25															
Ser	Ile	Leu	Leu	Asn	Pro	His	Thr	Gly	Gln	Pro	Phe	Val	Asn	Pro	Asp
480						485					490				
gga	act	cct	gca	ata	tac	aac	cca	ccc	acc	agt	cag	cag	ccc	ctg	cga
Gly	Thr	Pro	Ala	Ile	Tyr	Asn	Pro	Pro	Thr	Ser	Gln	Gln	Pro	Leu	Arg
495					500					505					510
agc	gcc	atg	gtg	ggg	cag	tcc	caa	cag	cag	cca	cca	cag	cag	cag	ccc
Ser	Ala	Met	Val	Gly	Gln	Ser	Gln	Gln	Gln	Pro	Pro	Gln	Gln	Gln	Pro
				515						520					525
tcc	ccg	cag	ccc	caa	cag	cag	gtc	cag	cca	ccg	cag	cca	cag	atg	gca
Ser	Pro	Gln	Pro	Gln	Gln	Gln	Val	Gln	Pro	Pro	Gln	Pro	Gln	Met	Ala
				530						535					540
ggc	cct	ctg	gtc	act	cag	tct	gtc	cag	ggg	ctg	cag	gct	tcc	tcc	cag
Gly	Pro	Leu	Val	Thr	Gln	Ser	Val	Gln	Gly	Leu	Gln	Ala	Ser	Ser	Gln
				545						550					555
tca	gtg	caa	tat	cca	gca	gtc	tct	ttt	cct	ccc	cag	cac	ctc	cta	cct
Ser	Val	Gln	Tyr	Pro	Ala	Val	Ser	Phe	Pro	Pro	Gln	His	Leu	Leu	Pro
				560											570
gtg	tct	cca	acg	cag	cac	ttt	ccc	atg	aga	gat	gat	gtg	gca	aca	cag
Val	Ser	Pro	Thr	Gln	His	Phe	Pro	Met	Arg	Asp	Asp	Val	Ala	Thr	Gln
				575											590
ttt	ggc	cag	atg	acc	ctg	agc	cgg	cag	tcc	tcg	ggg	gag	act	cct	gaa
Phe	Gly	Gln	Met	Thr	Leu	Ser	Arg	Gln	Ser	Ser	Gly	Glu	Thr	Pro	Glu
				595											605
ccc	cca	tca	ggt	cct	gtc	tac	cca	tcc	tcc	ctt	atg	cca	cag	ccg	gcc
Pro	Pro	Ser	Gly	Pro	Val	Tyr	Pro	Ser	Ser	Leu	Met	Pro	Gln	Pro	Ala
				610											620
cag	cag	ccc	agc	tat	gta	atc	gcc	tct	aca	ggc	cag	cag	ctt	cct	aca
Gln	Gln	Pro	Ser	Tyr	Val	Ile	Ala	Ser	Thr	Gly	Gln	Gln	Leu	Pro	Thr
				625											635
gga	gga	ttc	tca	ggc	tct	ggc	cct	ccc	atc	tcc	cag	cag	gtc	ctc	cag
Gly	Gly	Phe	Ser	Gly	Ser	Gly	Pro	Pro	Ile	Ser	Gln	Gln	Val	Leu	Gln
				640											650
ccc	cct	ccc	tca	cca	cag	gga	ttt	gtg	caa	cag	cct	ccg	cct	gca	cag
Pro	Pro	Pro	Ser	Pro	Gln	Gly	Phe	Val	Gln	Pro	Pro	Pro	Pro	Ala	Gln
				655											670
atg	cct	gta	tat	tat	tac	cca	tct	ggt	cag	tac	cct	acc	tca	acc	acg
Met	Pro	Val	Tyr	Tyr	Tyr	Pro	Ser	Gly	Gln	Tyr	Pro	Thr	Ser	Thr	Thr
				675											685
caa	cag	tac	cgg	ccc	atg	gcc	ccg	gtt	cag	tac	aac	gct	cag	agg	agt
Gln	Gln	Tyr	Arg	Pro	Met	Ala	Pro	Val	Gln	Tyr	Asn	Ala	Gln	Arg	Ser
				690											700
caa	cag	atg	cca	cag	gca	gca	cag	caa	gca	ggt	tac	cag	cca	gtc	ttg
Gln	Gln	Met	Pro	Gln	Ala	Ala	Gln	Gln	Ala	Gly	Tyr	Gln	Pro	Val	Leu
				705											715
tct	ggt	caa	cag	gga	ttc	caa	ggc	cta	ata	gga	gtg	cag	cag	cca	cct
Ser	Gly	Gln	Gln	Gly	Phe	Gln	Gly	Leu	Ile	Gly	Val	Gln	Gln	Pro	Pro
				720											730
cag	agt	cag	aac	gtg	ata	aat	aac	caa	caa	gga	act	ccg	gtg	caa	agc
Gln	Ser	Gln	Asn	Val	Ile	Asn	Asn	Gln	Gln	Gly	Thr	Pro	Val	Gln	Ser
				735											750
gtg	atg	gtt	tcc	tac	cca	aca	atg	tct	tct	tat	cag	gtg	cca	atg	acc
Val	Met	Val	Ser	Tyr	Pro	Thr	Met	Ser	Ser	Tyr	Gln	Val	Pro	Met	Thr
				755											765
cag	ggt	tct	caa	gga	ctg	ccc	cag	cag	tca	tac	caa	cag	cca	atc	atg
Gln	Gly	Ser	Gln	Gly	Leu	Pro	Gln	Gln	Ser	Tyr	Gln	Gln	Pro	Ile	Met
				770											780
cta	cct	aac	cag	gca	ggt	caa	ggg	tca	ctc	cca	gcc	act	gga	atg	cct
Leu	Pro	Asn	Gln	Ala	Gly	Gln	Gly	Ser	Leu	Pro	Ala	Thr	Gly	Met	Pro
				785											795

16U 100 PCT.ST25
 gtt tac tgt aat gtc aca ccg ccc acc cct cag aac aac ctt agg ctg 2691
 Val Tyr Cys Asn Val Thr Pro Pro Thr Pro Gln Asn Asn Leu Arg Leu
 800 805 810
 att ggc cca cac tgc ccc tcc agc act gtc cca gtg atg tca gct agc 2739
 Ile Gly Pro His Cys Pro Ser Ser Thr Val Pro Val Met Ser Ala Ser
 815 820 825 830
 tgc aga aca aac tgt gca agt atg agc aat gct ggt tgg cag gtc aaa 2787
 Cys Arg Thr Asn Cys Ala Ser Met Ser Asn Ala Gly Trp Gln Val Lys
 835 840 845
 ttc tga gagctctggc tgtggtacat ttcttcagat atttctcatg gcctttgatg 2843
 Phe
 gaagaggaac aagggtgggaa aactggctga ggacttaagt attcactcaa cactcaaatg 2903
 attgctgctg gtattctgta aaaaataaac aaagactaat atacacgtta gctgggtaatt 2963
 ggtgcatatt tctgtcatgt ctgctaggta tgcctttata gcttagctag tgacatgaat 3023
 tcatacaagg aagattttct cctaccactg aataccactg tgtagattat aatatcccta 3083
 atttggatta gttttgtact ttgtgttgag tttgtgatgc taaaagtatt taaaaattat 3143
 atactaaatc acattgtacc aaagctgtaa tggaaaagca aagaagaatt gatgaattga 3203
 aggaataatt tatatacatt atagagtttt cttttttaat ggatatatac tgtattgtag 3263
 tgtttaatca aaataaaact atttgacctt atggaggaag gtcattgttt taccacaaa 3323
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 3369

<210> 44
 <211> 847
 <212> PRT
 <213> Homo sapiens

<400> 44

Met Ser Glu Gln Gly Asp Leu Asn Gln Ala Ile Ala Glu Glu Gly Gly
 1 5 10 15
 Thr Glu Gln Glu Thr Ala Thr Pro Glu Asn Gly Ile Val Lys Ser Glu
 20 25 30
 Ser Leu Asp Glu Glu Glu Lys Leu Glu Leu Gln Arg Arg Leu Glu Ala
 35 40 45
 Gln Asn Gln Glu Arg Arg Lys Ser Lys Ser Gly Ala Gly Lys Gly Lys
 50 55 60
 Leu Thr Arg Ser Leu Ala Val Cys Glu Glu Ser Ser Ala Arg Pro Gly
 65 70 75 80
 Gly Glu Ser Leu Gln Asp Gln Glu Ser Ile His Leu Gln Leu Ser Ser
 85 90 95
 Phe Ser Ser Leu Gln Glu Glu Asp Lys Ser Arg Lys Asp Asp Ser Glu
 100 105 110
 Arg Glu Lys Glu Lys Asp Lys Asn Lys Asp Lys Thr Ser Glu Lys Pro
 115 120 125
 Lys Ile Arg Met Leu Ser Lys Asp Cys Ser Gln Glu Tyr Thr Asp Ser
 130 135 140
 Thr Gly Ile Asp Leu His Glu Phe Leu Ile Asn Thr Leu Lys Asn Asn

Page 62

16U 100 PCT.ST25

Ile Leu Leu Pro Leu Glu Ala Ala Thr Gly Ile Pro Pro Gly Ser Ile
465 470 475 480

Leu Leu Asn Pro His Thr Gly Gln Pro Phe Val Asn Pro Asp Gly Thr
485 490 495

Pro Ala Ile Tyr Asn Pro Pro Thr Ser Gln Gln Pro Leu Arg Ser Ala
500 505 510

Met Val Gly Gln Ser Gln Gln Gln Pro Pro Gln Gln Gln Pro Ser Pro
515 520 525

Gln Pro Gln Gln Gln Val Gln Pro Pro Gln Pro Gln Met Ala Gly Pro
530 535 540

Leu Val Thr Gln Ser Val Gln Gly Leu Gln Ala Ser Ser Gln Ser Val
545 550 555 560

Gln Tyr Pro Ala Val Ser Phe Pro Pro Gln His Leu Leu Pro Val Ser
565 570 575

Pro Thr Gln His Phe Pro Met Arg Asp Asp Val Ala Thr Gln Phe Gly
580 585 590

Gln Met Thr Leu Ser Arg Gln Ser Ser Gly Glu Thr Pro Glu Pro Pro
595 600 605

Ser Gly Pro Val Tyr Pro Ser Ser Leu Met Pro Gln Pro Ala Gln Gln
610 615 620

Pro Ser Tyr Val Ile Ala Ser Thr Gly Gln Gln Leu Pro Thr Gly Gly
625 630 635 640

Phe Ser Gly Ser Gly Pro Pro Ile Ser Gln Gln Val Leu Gln Pro Pro
645 650 655

Pro Ser Pro Gln Gly Phe Val Gln Gln Pro Pro Pro Ala Gln Met Pro
660 665 670

Val Tyr Tyr Tyr Pro Ser Gly Gln Tyr Pro Thr Ser Thr Thr Gln Gln
675 680 685

Tyr Arg Pro Met Ala Pro Val Gln Tyr Asn Ala Gln Arg Ser Gln Gln
690 695 700

Met Pro Gln Ala Ala Gln Gln Ala Gly Tyr Gln Pro Val Leu Ser Gly
705 710 715 720

Gln Gln Gly Phe Gln Gly Leu Ile Gly Val Gln Gln Pro Pro Gln Ser
725 730 735

Gln Asn Val Ile Asn Asn Gln Gln Gly Thr Pro Val Gln Ser Val Met
740 745 750

Val Ser Tyr Pro Thr Met Ser Ser Tyr Gln Val Pro Met Thr Gln Gly
755 760 765

Ser Gln Gly Leu Pro Gln Gln Ser Tyr Gln Gln Pro Ile Met Leu Pro
770 775 780

16U 100 PCT.ST25

Asn Gln Ala Gly Gln Gly Ser Leu Pro Ala Thr Gly Met Pro Val Tyr
785 790 795 800

Cys Asn Val Thr Pro Pro Thr Pro Gln Asn Asn Leu Arg Leu Ile Gly
805 810 815

Pro His Cys Pro Ser Ser Thr Val Pro Val Met Ser Ala Ser Cys Arg
820 825 830

Thr Asn Cys Ala Ser Met Ser Asn Ala Gly Trp Gln Val Lys Phe
835 840 845

<210> 45
<211> 3374
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (329)..(2812)
<223>

<400> 45
gtctattttt aatgctattt aatgaaggag cgagcgctc actcagcaat aaaagaagca 60
tgagggaaga cagagcagtg catggttatg gatactggac aaggatattt ggaaagggtg 120
acgatgtgtc acactgtgta agggaaatcgc atggagatgg gcattccgaa ctgttaatgg 180
ggacatggga ctccagtgtg ctctgatcac ttgtgtggat tttcctggcg tagaacgaca 240
gaagccgcta gtaagtcgcc aagacctaca gcaggaattc tgcaccaaag ggcataaaat 300
cttgttattt taatttgcac ctggggaga atg tct gag caa gga gac ctg aat 352
Met Ser Glu Gln Gly Asp Leu Asn
1 5
cag gca ata gca gag gaa gga ggg act gag cag gag acg gcc act cca 400
Gln Ala Ile Ala Glu Glu Gly Gly Thr Glu Gln Glu Thr Ala Thr Pro
10 15 20
gag aac ggc att gtt aaa tca gaa agt ctg gat gaa gag gag aaa ctg 448
Glu Asn Gly Ile Val Lys Ser Glu Ser Leu Asp Glu Glu Glu Lys Leu
25 30 35 40
gaa ctg cag agg cgg ctg gag gct cag aat caa gaa aga aga aaa tcc 496
Glu Leu Gln Arg Arg Leu Glu Ala Gln Asn Gln Glu Arg Arg Lys Ser
45 50 55
aag tca gga gca gga aaa ggt aaa ctg act cgc agt ctt gct gtc tgt 544
Lys Ser Gly Ala Gly Lys Gly Lys Leu Thr Arg Ser Leu Ala Val Cys
60 65 70
gag gaa tct tct gcc aga cca gga ggt gaa agt ctt cag gat cag gaa 592
Glu Glu Ser Ser Ala Arg Pro Gly Gly Glu Ser Leu Gln Asp Gln Glu
75 80 85
tca att cat tta cag ctt tcc agt ttt tcc agc ctg caa gag gag gat 640
Ser Ile His Leu Gln Leu Ser Ser Phe Ser Ser Leu Gln Glu Glu Asp
90 95 100
aaa tct agg aaa gat gac tct gaa aga gaa aaa gaa aag gat aaa aac 688
Lys Ser Arg Lys Asp Ser Glu Arg Glu Lys Glu Lys Asp Lys Asn
105 110 115 120
aaa gat aaa acc tct gaa aaa ccc aag atc aga atg tta tca aaa gat 736
Lys Asp Lys Thr Ser Glu Lys Pro Lys Ile Arg Met Leu Ser Lys Asp
125 130 135
tgc agc caa gaa tac acg gat tct aca ggc ata gac tta cac gag ttt 784
Cys Ser Gln Glu Tyr Thr Asp Ser Thr Gly Ile Asp Leu His Glu Phe
140 145 150
ctg att aac aca tta aag aat aat tcc agg gac agg atg ata ctt ttg 832

16U 100 PCT.ST25

Leu	Ile	Asn	Thr	Leu	Lys	Asn	Asn	Ser	Arg	Asp	Arg	Met	Ile	Leu	Leu	
		155					160					165				
aaa	atg	gag	cag	gaa	att	att	gat	ttc	att	gct	gac	aac	aat	aat	cat	880
Lys	Met	Glu	Gln	Glu	Ile	Ile	Asp	Phe	Ile	Ala	Asp	Asn	Asn	Asn	His	
	170					175					180					
tat	aaa	aag	ttc	cct	cag	atg	tca	tcg	tat	cag	agg	atg	ctt	gtc	cat	928
Tyr	Lys	Lys	Phe	Pro	Gln	Met	Ser	Ser	Tyr	Gln	Arg	Met	Leu	Val	His	
	185				190					195				200		
cga	gtg	gca	gct	tat	ttt	gga	ttg	gat	cac	aat	gtg	gat	caa	aca	gga	976
Arg	Val	Ala	Ala	Tyr	Phe	Gly	Leu	Asp	His	Asn	Val	Asp	Gln	Thr	Gly	
				205					210					215		
aaa	tct	gtt	atc	atc	aac	aag	acc	agc	agc	acc	aga	ata	cca	gag	caa	1024
Lys	Ser	Val	Ile	Ile	Asn	Lys	Thr	Ser	Ser	Thr	Arg	Ile	Pro	Glu	Gln	
			220					225					230			
agg	ttt	tgt	gaa	cat	tta	aaa	gat	gaa	aaa	ggg	gaa	gaa	tcc	cag	aag	1072
Arg	Phe	Cys	Glu	His	Leu	Lys	Asp	Glu	Lys	Gly	Glu	Ser	Gln	Lys		
		235					240					245				
cgg	ttt	atc	ttg	aag	cga	gat	aac	tct	agt	att	gat	aaa	gaa	gac	aat	1120
Arg	Phe	Ile	Leu	Lys	Arg	Asp	Asn	Ser	Ser	Ile	Asp	Lys	Glu	Asp	Asn	
		250				255					260					
cag	caa	aac	aga	atg	cat	cca	ttt	aga	gat	gac	aga	cga	agt	aaa	tca	1168
Gln	Gln	Asn	Arg	Met	His	Pro	Phe	Arg	Asp	Asp	Arg	Arg	Ser	Lys	Ser	
				270						275					280	
att	gaa	gag	aga	gaa	gag	gaa	tat	cag	aga	gtg	agg	gag	aga	ata	ttt	1216
Ile	Glu	Glu	Arg	Glu	Glu	Glu	Tyr	Gln	Arg	Val	Arg	Glu	Arg	Ile	Phe	
				285					290					295		
gca	cac	gat	tca	gtt	tgc	tcc	cag	gaa	agc	ctt	ttt	gtg	gaa	aac	agg	1264
Ala	His	Asp	Ser	Val	Cys	Ser	Gln	Ser	Leu	Phe	Val	Glu	Asn	Arg		
				300				305				310				
ggc	aac	aga	gat	ggc	tca	ggg	aga	aca	tct	ggg	agt	cga	cag	agc	agc	1312
Gly	Asn	Arg	Arg	Gly	Ser	Gly	Arg	Thr	Ser	Gly	Ser	Gln	Ser	Ser		
		315					320					325				
tca	gaa	aat	gaa	ctc	aag	tgg	tct	gac	cac	caa	agg	gcc	tgg	agc	agc	1360
Ser	Glu	Asn	Glu	Leu	Lys	Trp	Ser	Asp	His	Gln	Arg	Ala	Trp	Ser	Ser	
		330				335					340					
aca	gac	tcc	gac	agt	tcc	aac	cgc	aat	cta	aag	ccc	gcc	atg	acc	aag	1408
Thr	Asp	Ser	Asp	Ser	Ser	Asn	Arg	Asn	Leu	Lys	Pro	Ala	Met	Thr	Lys	
		345			350					355					360	
acg	gcg	agt	ttt	ggg	ggc	atc	acg	gtg	ctg	acc	agg	ggg	gac	agc	act	1456
Thr	Ala	Ser	Phe	Gly	Gly	Ile	Thr	Val	Leu	Thr	Arg	Gly	Asp	Ser	Thr	
				365					370					375		
tcc	agt	act	agg	agt	acc	ggg	aag	ctg	tcc	aaa	gca	ggg	tcc	gag	tct	1504
Ser	Ser	Thr	Arg	Ser	Thr	Gly	Lys	Leu	Ser	Lys	Ala	Gly	Ser	Glu	Ser	
			380					385					390			
tcc	agc	agt	gca	ggc	tcc	tca	gga	tcg	ctg	tcc	cgc	acc	cat	cca	cct	1552
Ser	Ser	Ser	Ala	Gly	Ser	Ser	Gly	Ser	Leu	Ser	Arg	Thr	His	Pro	Pro	
			395				400					405				
ctc	cag	agc	aca	ccc	cta	gtc	tca	ggg	gtg	gca	gct	ggc	tct	cca	ggc	1600
Leu	Gln	Ser	Thr	Pro	Leu	Val	Ser	Gly	Val	Ala	Ala	Gly	Ser	Pro	Gly	
		410				415					420					
tgt	gtg	cct	tat	cca	gag	aat	gga	ata	ggg	ggc	cag	gtt	gct	ccc	agc	1648
Cys	Val	Pro	Tyr	Pro	Glu	Asn	Gly	Ile	Gly	Gly	Gln	Val	Ala	Pro	Ser	
		425			430				435					440		
agc	acc	agc	tac	atc	ctc	ctt	cca	ctt	gaa	gct	gca	aca	ggc	atc	ccg	1696
Ser	Thr	Ser	Tyr	Ile	Leu	Leu	Pro	Leu	Glu	Ala	Ala	Thr	Gly	Ile	Pro	
				445					450					455		
cct	gga	agc	atc	ctt	ctt	aat	cca	cac	aca	ggc	cag	ccc	ttt	gtg	aat	1744
Pro	Gly	Ser	Ile	Leu	Leu	Asn	Pro	His	Thr	Gly	Gln	Pro	Phe	Val	Asn	
			460				465						470			

16U 100 PCT.ST25																
ccc Pro	gat Asp	gga Gly 475	act Thr	cct Pro	gca Ala	ata Ile	tac Tyr 480	aac Asn	cca Pro	ccc Pro	acc Thr	agt Ser 485	cag Gln	cag Gln	ccc Pro	1792
ctg Leu	cga Arg 490	agc Ser	gcc Ala	atg Met	gtg Val	ggg Gly 495	cag Gln	tcc Ser	caa Gln	cag Gln	cag Gln 500	ccg Pro	cca Pro	cag Gln	cag Gln	1840
cag Gln 505	ccc Pro	tcc Ser	ccg Pro	cag Gln	ccc Pro 510	caa Gln	cag Gln	cag Gln	gtc Val	cag Gln 515	cca Pro	ccg Pro	cag Gln	cca Pro	cag Gln 520	1888
atg Met	gca Ala	ggc Gly	cct Pro	ctg Leu 525	gtc Val	act Thr	cag Gln	tct Ser	gtc Val 530	cag Gln	ggg Gly	ctg Leu	cag Gln	gct Ala 535	tcc Ser	1936
tcc Ser	cag Gln	tca Ser	gtg Val 540	caa Gln	tat Tyr	ccg Pro	gca Ala	gtc Val 545	tct Ser	ttt Phe	cct Pro	ccc Pro	cag Gln 550	cac His	ctc Leu	1984
cta Leu	cct Pro	gtg Val 555	tct Ser	cca Pro	acg Thr	cag Gln	cac His 560	ttt Phe	ccc Pro	atg Met	aga Arg	gat Asp 565	gat Asp	gtg Val	gca Ala	2032
aca Thr	cag Gln 570	ttt Phe	ggc Gly	cag Gln	atg Met	acc Thr 575	ctg Leu	agc Ser	cgg Arg	cag Gln	tcc Ser 580	tcg Ser	ggg Gly	gag Glu	act Thr	2080
cct Pro 585	gaa Glu	ccc Pro	cca Pro	tca Ser	ggt Gly 590	cct Pro	gtc Val	tac Tyr	cca Pro	tcc Ser 595	tcc Ser	ctt Leu	atg Met	cca Pro	cag Gln 600	2128
ccg Pro	gcc Ala	cag Gln	cag Gln	ccc Pro 605	agc Ser	tat Tyr	gta Val	atc Ile	gcc Ala 610	tct Ser	aca Thr	ggc Gly	cag Gln	cag Gln 615	ctt Leu	2176
cct Pro	aca Thr	gga Gly	gga Gly 620	ttc Phe	tca Ser	ggc Gly	tct Ser	ggc Gly 625	cct Pro	ccc Pro	atc Ile	tcc Ser	cag Gln 630	cag Gln	gtc Val	2224
ctc Leu	cag Gln	ccc Pro 635	cct Pro	ccc Pro	tca Ser	cca Pro	cag Gln 640	gga Gly	ttt Phe	gtg Val	caa Gln	cag Gln 645	cct Pro	ccg Pro	cct Pro	2272
gca Ala	cag Gln 650	atg Met	cct Pro	gta Val	tat Tyr	tat Tyr 655	tac Tyr	cca Pro	tct Ser	ggt Gly	cag Gln 660	tac Tyr	cct Pro	acc Thr	tca Ser	2320
acc Thr 665	acg Thr	caa Gln	cag Gln	tac Tyr	cgg Arg 670	ccc Pro	atg Met	gcc Ala	ccg Pro	gtt Val 675	cag Gln	tac Tyr	aac Asn	gct Ala	cag Gln 680	2368
agg Arg	agt Ser	caa Gln	cag Gln	atg Met 685	cca Pro	cag Gln	gca Ala	gca Ala	cag Gln 690	caa Gln	gca Ala	ggt Gly	tac Tyr	cag Gln 695	cca Pro	2416
gtc Val	ttg Leu	tct Ser	ggt Gly 700	caa Gln	cag Gln	gga Gly	ttc Phe	caa Gln 705	ggc Gly	cta Leu	ata Ile	gga Gly 710	gtg Val	cag Gln	cag Gln	2464
cca Pro	cct Pro	cag Gln 715	agt Ser	cag Gln	aac Asn	gtg Val 720	ata Ile	aat Asn	aac Asn	caa Gln	caa Gln 725	gga Gly 725	act Thr	ccg Pro	gtg Val	2512
caa Gln	agc Ser	gtg Val 730	atg Met	gtt Val	tcc Ser	tac Tyr 735	cca Pro	aca Thr	atg Met	tct Ser	tct Ser 740	tat Tyr	cag Gln	gtg Val	cca Pro	2560
atg Met 745	acc Thr	cag Gln	ggt Gly	tct Ser	caa Gln 750	gga Gly	ctg Leu	ccc Pro	cag Gln	cag Gln 755	tca Ser	tac Tyr	caa Gln	cag Gln	cca Pro 760	2608
atc Ile	atg Met	cta Leu	cct Pro	aac Asn 765	cag Gln	gca Ala	ggt Gly	caa Gln	ggg Gly 770	tca Ser	ctc Leu	cca Pro	gcc Ala	act Thr 775	gga Gly	2656
atg Met	cct Pro	gtt Val 780	tac Tyr	tgt Cys	aat Asn	gtc Val	aca Thr	ccg Pro 785	ccc Pro	acc Thr	cct Pro	cag Gln	aac Asn 790	aac Asn	ctt Leu	2704

16U 100 PCT.ST25

```

agg ctg att ggc cca cac tgc ccc tcc agc act gtc cca gtg atg tca 2752
Arg Leu Ile Gly Pro His Cys Pro Ser Ser Thr Val Pro Val Met Ser
      795                      800                      805

gct agc tgc aga aca aac tgt gca agt atg agc aat gct ggt tgg cag 2800
Ala Ser Cys Arg Thr Asn Cys Ala Ser Met Ser Asn Ala Gly Trp Gln
      810                      815                      820

gtc aaa ttc tga gagctctggc tgtggtacat ttcttcagat atttctcatg 2852
Val Lys Phe
825

gcctttgatg gaagaggaac aaggtgggaa aactggctga ggacttaagt attcactcaa 2912
cactcaaatg attgctgctg gtattctgtgta aaaaaataaac aaagactaat atacacgtta 2972
gctggttaat ggtgcatatt tctgtcatgt ctgctaggta tgcctttata gcttagctag 3032
tgacatgaat tcatcaagggt aagattttct cctaccactg aataccactg tgtagattat 3092
aatatcccta atttggatta gttttgtact ttgtgttgag tttgtgatgc taaaagtatt 3152
taaaaattat atactaaatc acattgtacc aaagctgtaa tggaaaagca aagaagaatt 3212
gatgaattga aggaataatt tatatacatt atagagtttt cttttttaat ggatatatac 3272
tgtattgtag tgtttaatca aaataaaact atttgacctt atggaggaag gtcattgttt 3332
taaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 3374

```

<210> 46
 <211> 827
 <212> PRT
 <213> Homo sapiens

<400> 46

```

Met Ser Glu Gln Gly Asp Leu Asn Gln Ala Ile Ala Glu Glu Gly Gly
1      5      10

Thr Glu Gln Glu Thr Ala Thr Pro Glu Asn Gly Ile Val Lys Ser Glu
20     25     30

Ser Leu Asp Glu Glu Glu Lys Leu Glu Leu Gln Arg Arg Leu Glu Ala
35     40     45

Gln Asn Gln Glu Arg Arg Lys Ser Lys Ser Gly Ala Gly Lys Gly Lys
50     55     60

Leu Thr Arg Ser Leu Ala Val Cys Glu Glu Ser Ser Ala Arg Pro Gly
65     70     75     80

Gly Glu Ser Leu Gln Asp Gln Glu Ser Ile His Leu Gln Leu Ser Ser
85     90     95

Phe Ser Ser Leu Gln Glu Glu Asp Lys Ser Arg Lys Asp Asp Ser Glu
100    105    110

Arg Glu Lys Glu Lys Asp Lys Asn Lys Asp Lys Thr Ser Glu Lys Pro
115    120    125

Lys Ile Arg Met Leu Ser Lys Asp Cys Ser Gln Glu Tyr Thr Asp Ser
130    135    140

Thr Gly Ile Asp Leu His Glu Phe Leu Ile Asn Thr Leu Lys Asn Asn
145    150    155    160

```

160 100 PCT.ST25
 Ser Arg Asp Arg Met Ile Leu Leu Lys Met Glu Gln Glu Ile Ile Asp
 165 170 175
 Phe Ile Ala Asp Asn Asn Asn His Tyr Lys Lys Phe Pro Gln Met Ser
 180 185 190
 Ser Tyr Gln Arg Met Leu Val His Arg Val Ala Ala Tyr Phe Gly Leu
 195 200 205
 Asp His Asn Val Asp Gln Thr Gly Lys Ser Val Ile Ile Asn Lys Thr
 210 215 220
 Ser Ser Thr Arg Ile Pro Glu Gln Arg Phe Cys Glu His Leu Lys Asp
 225 230 235 240
 Glu Lys Gly Glu Glu Ser Gln Lys Arg Phe Ile Leu Lys Arg Asp Asn
 245 250 255
 Ser Ser Ile Asp Lys Glu Asp Asn Gln Gln Asn Arg Met His Pro Phe
 260 265 270
 Arg Asp Asp Arg Arg Ser Lys Ser Ile Glu Glu Arg Glu Glu Glu Tyr
 275 280 285
 Gln Arg Val Arg Glu Arg Ile Phe Ala His Asp Ser Val Cys Ser Gln
 290 295 300
 Glu Ser Leu Phe Val Glu Asn Arg Gly Asn Arg Asp Gly Ser Gly Arg
 305 310 315 320
 Thr Ser Gly Ser Arg Gln Ser Ser Ser Glu Asn Glu Leu Lys Trp Ser
 325 330 335
 Asp His Gln Arg Ala Trp Ser Ser Thr Asp Ser Asp Ser Ser Asn Arg
 340 345 350
 Asn Leu Lys Pro Ala Met Thr Lys Thr Ala Ser Phe Gly Gly Ile Thr
 355 360 365
 Val Leu Thr Arg Gly Asp Ser Thr Ser Ser Thr Arg Ser Thr Gly Lys
 370 375 380
 Leu Ser Lys Ala Gly Ser Glu Ser Ser Ser Ser Ala Gly Ser Ser Gly
 385 390 395 400
 Ser Leu Ser Arg Thr His Pro Pro Leu Gln Ser Thr Pro Leu Val Ser
 405 410 415
 Gly Val Ala Ala Gly Ser Pro Gly Cys Val Pro Tyr Pro Glu Asn Gly
 420 425 430
 Ile Gly Gly Gln Val Ala Pro Ser Ser Thr Ser Tyr Ile Leu Leu Pro
 435 440 445
 Leu Glu Ala Ala Thr Gly Ile Pro Pro Gly Ser Ile Leu Leu Asn Pro
 450 455 460
 His Thr Gly Gln Pro Phe Val Asn Pro Asp Gly Thr Pro Ala Ile Tyr
 465 470 475 480

16U 100 PCT.ST25

Asn Pro Pro Thr Ser Gln Gln Pro Leu Arg Ser Ala Met Val Gly Gln
 485 490 495
 Ser Gln Gln Gln Pro Pro Gln Gln Gln Pro Ser Pro Gln Pro Gln Gln
 500 505 510
 Gln Val Gln Pro Pro Gln Pro Gln Met Ala Gly Pro Leu Val Thr Gln
 515 520 525
 Ser Val Gln Gly Leu Gln Ala Ser Ser Gln Ser Val Gln Tyr Pro Ala
 530 535 540
 Val Ser Phe Pro Pro Gln His Leu Leu Pro Val Ser Pro Thr Gln His
 545 550 555 560
 Phe Pro Met Arg Asp Asp Val Ala Thr Gln Phe Gly Gln Met Thr Leu
 565 570 575
 Ser Arg Gln Ser Ser Gly Glu Thr Pro Glu Pro Pro Ser Gly Pro Val
 580 585 590
 Tyr Pro Ser Ser Leu Met Pro Gln Pro Ala Gln Gln Pro Ser Tyr Val
 595 600 605
 Ile Ala Ser Thr Gly Gln Gln Leu Pro Thr Gly Gly Phe Ser Gly Ser
 610 615 620
 Gly Pro Pro Ile Ser Gln Gln Val Leu Gln Pro Pro Pro Ser Pro Gln
 625 630 635 640
 Gly Phe Val Gln Gln Pro Pro Pro Ala Gln Met Pro Val Tyr Tyr Tyr
 645 650 655
 Pro Ser Gly Gln Tyr Pro Thr Ser Thr Thr Gln Gln Tyr Arg Pro Met
 660 665 670
 Ala Pro Val Gln Tyr Asn Ala Gln Arg Ser Gln Gln Met Pro Gln Ala
 675 680 685
 Ala Gln Gln Ala Gly Tyr Gln Pro Val Leu Ser Gly Gln Gln Gly Phe
 690 695 700
 Gln Gly Leu Ile Gly Val Gln Gln Pro Pro Gln Ser Gln Asn Val Ile
 705 710 715 720
 Asn Asn Gln Gln Gly Thr Pro Val Gln Ser Val Met Val Ser Tyr Pro
 725 730 735
 Thr Met Ser Ser Tyr Gln Val Pro Met Thr Gln Gly Ser Gln Gly Leu
 740 745 750
 Pro Gln Gln Ser Tyr Gln Gln Pro Ile Met Leu Pro Asn Gln Ala Gly
 755 760 765
 Gln Gly Ser Leu Pro Ala Thr Gly Met Pro Val Tyr Cys Asn Val Thr
 770 775 780
 Pro Pro Thr Pro Gln Asn Asn Leu Arg Leu Ile Gly Pro His Cys Pro
 785 790 795 800

16U 100 PCT.ST25

Ser Ser Thr Val Pro Val Met Ser Ala Ser Cys Arg Thr Asn Cys Ala
805 810 815

Ser Met Ser Asn Ala Gly Trp Gln Val Lys Phe
820 825

<210> 47
<211> 3332
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (329)..(2770)
<223>

<400> 47
gtctattttt aatgctattt aatgaaggag cgagcgcctc actcagcaat aaaagaagca 60
tgagggaaga cagagcagtg catggttatg gatactggac aaggatattt ggaaagggtg 120
acgatgtgtc acactgtgta agggaatcgc atggagatgg gcattccgaa ctgttaatgg 180
ggacatggga ctccagtgtg ctctgatcac ttgtgtggat tttcctggcg tagaacgaca 240
gaagccgcta gtaagtcgcc aagacctaca gcaggaattc tgcaccaaag ggcataaaat 300
cttgttattt taatttgcac ctgggaga atg tct gag caa gga gac ctg aat 352
Met Ser Glu Gln Gly Asp Leu Asn
1 5
cag gca ata gca gag gaa gga ggg act gag cag gag acg gcc act cca 400
Gln Ala Ile Ala Glu Gly Thr Glu Gln Glu Thr Ala Thr Pro
10 15 20
gag aac ggc att gtt aaa tca gaa agt ctg gat gaa gag gag aaa ctg 448
Glu Asn Gly Ile Val Lys Ser Glu Ser Leu Asp Glu Glu Glu Lys Leu
25 30 35 40
gaa ctg cag agg cgg ctg gag gct cag aat caa gaa aga aga aaa tcc 496
Glu Leu Gln Arg Arg Leu Glu Ala Gln Asn Gln Glu Arg Arg Lys Ser
45 50 55
aag tca gga gca gga aaa ggt aaa ctg act cgc agt ctt gct gtc tgt 544
Lys Ser Gly Ala Gly Lys Gly Lys Leu Thr Arg Ser Leu Ala Val Cys
60 65 70
gag gaa tct tct gcc aga cca gga ggt gaa agt ctt cag gat cag gaa 592
Glu Glu Ser Ser Ala Arg Pro Gly Gly Glu Ser Leu Gln Asp Gln Glu
75 80 85
tca att cat tta cag ctt tcc agt ttt tcc agc ctg caa gag gag gat 640
Ser Ile His Leu Gln Leu Ser Ser Phe Ser Ser Leu Gln Glu Glu Asp
90 95 100
aaa tct agg aaa gat gac tct gaa aga gaa aaa gaa aag gat aaa aac 688
Lys Ser Arg Lys Asp Asp Ser Glu Arg Glu Lys Glu Lys Asp Lys Asn
105 110 115 120
aaa gat aaa acc tct gaa aaa ccc aag atc aga atg tta tca aaa gat 736
Lys Asp Lys Thr Ser Glu Lys Pro Lys Ile Arg Met Leu Ser Lys Asp
125 130 135
tgc agc caa gaa tac acg gat tct aca ggc ata gac tta cac gag ttt 784
Cys Ser Gln Glu Tyr Thr Asp Ser Thr Gly Ile Asp Leu His Glu Phe
140 145 150
ctg att aac aca tta aag aat aat tcc agg gac agg atg ata ctt ttg 832
Leu Ile Asn Thr Leu Lys Asn Asn Ser Arg Asp Arg Met Ile Leu Leu
155 160 165
aaa atg gag cag gaa att att gat ttc att gct gac aac aat aat cat 880
Lys Met Glu Gln Glu Ile Ile Asp Phe Ile Ala Asp Asn Asn Asn His
170 175 180

160 100 PCT.ST25															
tat	aaa	aag	ttc	cct	cag	atg	tca	tcg	tat	cag	agg	atg	ctt	gtc	cat
Tyr	Lys	Lys	Phe	Pro	Gln	Met	Ser	Ser	Tyr	Gln	Arg	Met	Leu	Val	His
185					190					195					200
cga	gtg	gca	gct	tat	ttt	gga	ttg	gat	cac	aat	gtg	gat	caa	aca	gga
Arg	Val	Ala	Ala	Tyr	Phe	Gly	Leu	Asp	His	Asn	Val	Asp	Gln	Thr	Gly
				205					210					215	
aaa	tct	ggt	atc	aac	aag	acc	agc	agc	acc	aga	ata	cca	gag	caa	
Lys	Ser	Val	Ile	Asn	Lys	Thr	Ser	Ser	Thr	Arg	Ile	Pro	Glu	Gln	
			220				225					230			
agg	ttt	tgt	gaa	cat	tta	aaa	gat	gaa	aaa	ggt	gaa	gaa	tcc	cag	aag
Arg	Phe	Cys	Glu	His	Leu	Lys	Asp	Glu	Lys	Gly	Glu	Glu	Ser	Gln	Lys
		235					240					245			
cgg	ttt	atc	ttg	aag	cga	gat	aac	tct	agt	att	gat	aaa	gaa	gac	aat
Arg	Phe	Ile	Leu	Lys	Arg	Asp	Asn	Ser	Ser	Ile	Asp	Lys	Glu	Asp	Asn
	250					255					260				
cag	tca	ggt	tgc	tcc	cag	gaa	agc	ctt	ttt	gtg	gaa	aac	agt	agg	ctc
Gln	Ser	Val	Cys	Ser	Gln	Glu	Ser	Leu	Phe	Val	Glu	Asn	Ser	Arg	Leu
265					270					275					280
ttg	gaa	gac	agt	aac	ata	tgc	aat	gag	acc	tat	aag	aaa	aga	cag	ctc
Leu	Glu	Asp	Ser	Asn	Ile	Cys	Asn	Glu	Thr	Tyr	Lys	Lys	Arg	Gln	Leu
				285						290				295	
ttt	cgg	ggc	aac	aga	gat	ggc	tca	ggg	aga	aca	tct	ggg	agt	cga	cag
Phe	Arg	Gly	Asn	Arg	Asp	Gly	Ser	Gly	Arg	Thr	Ser	Gly	Ser	Arg	Gln
			300					305					310		
agc	agc	tca	gaa	aat	gaa	ctc	aag	tgg	tct	gac	cac	caa	agg	gcc	tgg
Ser	Ser	Ser	Glu	Asn	Glu	Leu	Lys	Trp	Ser	Asp	His	Gln	Arg	Ala	Trp
			315				320					325			
agc	agc	aca	gac	tcc	gac	agt	tcc	aac	cgc	aat	cta	aag	ccc	gcc	atg
Ser	Ser	Thr	Asp	Ser	Asp	Ser	Ser	Asn	Arg	Asn	Leu	Lys	Pro	Ala	Met
			330			335					340				
acc	aag	acg	gcg	agt	ttt	ggg	ggc	atc	acg	gtg	ctg	acc	agg	ggt	gac
Thr	Lys	Thr	Ala	Ser	Phe	Gly	Gly	Ile	Thr	Val	Leu	Thr	Arg	Gly	Asp
345					350					355					360
agc	act	tcc	agt	act	agg	agt	acc	ggg	aag	ctg	tcc	aaa	gca	ggt	tcc
Ser	Thr	Ser	Ser	Thr	Arg	Ser	Thr	Gly	Lys	Leu	Ser	Lys	Ala	Gly	Ser
				365					370					375	
gag	tct	tcc	agc	agt	gca	ggc	tcc	tca	gga	tcg	ctg	tcc	cgc	acc	cat
Glu	Ser	Ser	Ser	Ser	Ala	Gly	Ser	Ser	Gly	Ser	Leu	Ser	Arg	Thr	His
				380				385					390		
cca	cct	ctc	cag	agc	aca	ccc	cta	gtc	tca	ggt	gtg	gca	gct	ggc	tct
Pro	Pro	Leu	Gln	Ser	Thr	Pro	Leu	Val	Ser	Gly	Val	Ala	Ala	Gly	Ser
			395				400					405			
cca	ggc	tgt	gtg	cct	tat	cca	gag	aat	gga	ata	ggg	ggc	cag	ggt	gct
Pro	Gly	Cys	Val	Pro	Tyr	Pro	Glu	Asn	Gly	Ile	Gly	Gly	Gln	Val	Ala
			410			415					420				
ccc	agc	agc	acc	agc	tac	atc	ctc	ctt	cca	ctt	gaa	gct	gca	aca	ggc
Pro	Ser	Ser	Thr	Ser	Tyr	Ile	Leu	Leu	Pro	Leu	Glu	Ala	Ala	Thr	Gly
425					430					435					440
atc	ccg	cct	gga	agc	atc	ctt	ctt	aat	cca	cac	aca	ggc	cag	ccc	ttt
Ile	Pro	Pro	Gly	Ser	Ile	Leu	Leu	Asn	Pro	His	Thr	Gly	Gln	Pro	Phe
				445					450					455	
gtg	aat	ccc	gat	gga	act	cct	gca	ata	tac	aac	cca	ccc	acc	agt	cag
Val	Asn	Pro	Asp	Gly	Thr	Pro	Ala	Ile	Tyr	Asn	Pro	Pro	Thr	Ser	Gln
			460				465						470		
cag	ccc	ctg	cga	agc	gcc	atg	gtg	ggg	cag	tcc	caa	cag	cag	ccg	cca
Gln	Pro		Arg	Ser	Ala	Met	Val	Gly	Gln	Ser	Gln	Gln	Gln	Pro	Pro
		475					480					485			
cag	cag	cag	ccc	tcc	ccg	cag	ccc	caa	cag	cag	gtc	cag	cca	ccg	cag
Gln	Gln	Gln	Pro	Ser	Pro	Gln	Pro	Gln	Gln	Gln	Val	Gln	Pro	Pro	Gln
			490			495					500				

16U 100 PCT.ST25

cca cag atg gca ggc cct ctg gtc act cag tct gtc cag ggg ctg cag Pro Gln Met Ala Gly Pro Leu Val Thr Gln Ser Val Gln Gly Leu Gln 505 510 515 520	1888
gct tcc tcc cag tca gtg caa tat ccg gca gtc tct ttt cct ccc cag Ala Ser Ser Gln Ser Val Gln Tyr Pro Ala Val Ser Phe Pro Pro Gln 525 530 535	1936
cac ctc cta cct gtg tct cca acg cag cac ttt ccc atg aga gat gat His Leu Leu Pro Val Ser Pro Thr Gln His Phe Pro Met Arg Asp Asp 540 545 550	1984
gtg gca aca cag ttt ggc cag atg acc ctg agc cgg cag tcc tcg ggg Val Ala Thr Gln Phe Gly Gln Met Thr Leu Ser Arg Gln Ser Ser Gly 555 560 565	2032
gag act cct gaa ccc cca tca ggt cct gtc tac cca tcc tcc ctt atg Glu Thr Pro Glu Pro Pro Ser Gly Pro Val Tyr Pro Ser Ser Leu Met 570 575 580	2080
cca cag ccg gcc cag cag ccc agc tat gta atc gcc tct aca ggc cag Pro Gln Pro Ala Gln Gln Pro Ser Tyr Val Ile Ala Ser Thr Gly Gln 585 590 595 600	2128
cag ctt cct aca gga gga ttc tca ggc tct ggc cct ccc atc tcc cag Gln Leu Pro Thr Gly Gly Phe Ser Gly Ser Gly Pro Pro Ile Ser Gln 605 610 615	2176
cag gtc ctc cag ccc cct ccc tca cca cag gga tty gtg caa cag cct Gln Val Leu Gln Pro Pro Pro Ser Pro Gln Gly Phe Val Gln Gln Pro 620 625 630	2224
ccg cct gca cag atg cct gta tat tat tac cca tct ggt cag tac cct Pro Pro Ala Gln Met Pro Val Tyr Tyr Tyr Pro Ser Gly Gln Tyr Pro 635 640 645	2272
acc tca acc acg caa cag tac ccg ccc atg gcc ccg gtt cag tac aac Thr Ser Thr Thr Gln Gln Tyr Arg Pro Met Ala Pro Val Gln Tyr Asn 650 655 660	2320
gct cag agg agt caa cag atg cca cag gca gca cag caa gca ggt tac Ala Gln Arg Ser Gln Met Pro Gln Ala Ala Gln Gln Ala Gly Tyr 665 670 675 680	2368
cag cca gtc ttg tct ggt caa cag gga ttc caa ggc cta ata gga gtg Gln Pro Val Leu Ser Gly Gln Gln Gly Phe Gln Gly Leu Ile Gly Val 685 690 695	2416
cag cag cca cct cag agt cag aac gtg ata aat aac caa caa gga act Gln Gln Pro Pro Gln Ser Gln Asn Val Ile Asn Asn Gln Gln Gly Thr 700 705 710	2464
ccg gtg caa agc gtg atg gtt tcc tac cca aca atg tct tct tat cag Pro Val Gln Ser Val Met Val Ser Tyr Pro Thr Met Ser Ser Tyr Gln 715 720 725	2512
gtg cca atg acc cag ggt tct caa gga ctg ccc cag cag tca tac caa Val Pro Met Thr Gln Gly Ser Gln Gly Leu Pro Gln Gln Ser Tyr Gln 730 735 740	2560
cag cca atc atg cta cct aac cag gca ggt caa ggg tca ctc cca gcc Gln Pro Ile Met Leu Pro Asn Gln Ala Gly Gln Gly Ser Leu Pro Ala 745 750 755 760	2608
act gga atg cct gtt tac tgt aat gtc aca ccg ccc acc cct cag aac Thr Gly Met Pro Val Tyr Cys Asn Val Thr Pro Pro Thr Pro Gln Asn 765 770 775	2656
aac ctt agg ctg att ggc cca cac tgc ccc tcc agc act gtc cca gtg Asn Leu Arg Leu Ile Gly Pro His Cys Pro Ser Ser Thr Val Pro Val 780 785 790	2704
atg tca gct agc tgc aga aca aac tgt gca agt atg agc aat gct ggt Met Ser Ala Ser Cys Arg Thr Asn Cys Ala Ser Met Ser Asn Ala Gly 795 800 805	2752
tgg cag gtc aaa ttc tga gagctctggc tgtggtacat ttcttcagat Trp Gln Val Lys Phe	2800

810 16U 100 PCT.ST25

atttctcatg gcctttgatg gaagaggaac aaggtgggaa aactggctga ggacttaagt 2860
 attcactcaa cactcaaagtg attgctgctg gtattctgta aaaartaaac aaagactaat 2920
 atacacgtta gctgggtaat ggtgcatatt tctgtcatgt ctgctaggta tgcctttata 2980
 gcttagctag tgacatgaat tcatcaaggt aagattytct cctaccactg aataccactg 3040
 tgtagattat aatatcccta atttggatta gttttgtact ttgtgttgag tttgtgatgc 3100
 taaaagtatt taaaaattat atactaaatc acattgtacc aaagctgtaa tggaaaagca 3160
 aagaagaayt gatgaattga aggaataatt tatatacatt atagagtttt cttttttaat 3220
 ggatataatc tgtagttag tgtttaatca aaataaaact atttgacctt atggaggaag 3280
 gtcattgtttt taaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 3332

<210> 48
 <211> 813
 <212> PRT
 <213> Homo sapiens

<400> 48

Met Ser Glu Gln Gly Asp Leu Asn Gln Ala Ile Ala Glu Glu Gly Gly
 1 5 10 15

Thr Glu Gln Glu Thr Ala Thr Pro Glu Asn Gly Ile Val Lys Ser Glu
 20 25 30

Ser Leu Asp Glu Glu Glu Lys Leu Glu Leu Gln Arg Arg Leu Glu Ala
 35 40 45

Gln Asn Gln Glu Arg Arg Lys Ser Lys Ser Gly Ala Gly Lys Gly Lys
 50 55 60

Leu Thr Arg Ser Leu Ala Val Cys Glu Glu Ser Ser Ala Arg Pro Gly
 65 70 75 80

Gly Glu Ser Leu Gln Asp Gln Glu Ser Ile His Leu Gln Leu Ser Ser
 85 90 95

Phe Ser Ser Leu Gln Glu Glu Asp Lys Ser Arg Lys Asp Asp Ser Glu
 100 105 110

Arg Glu Lys Glu Lys Asp Lys Asn Lys Asp Lys Thr Ser Glu Lys Pro
 115 120 125

Lys Ile Arg Met Leu Ser Lys Asp Cys Ser Gln Glu Tyr Thr Asp Ser
 130 135 140

Thr Gly Ile Asp Leu His Glu Phe Leu Ile Asn Thr Leu Lys Asn Asn
 145 150 155 160

Ser Arg Asp Arg Met Ile Leu Leu Lys Met Glu Gln Glu Ile Ile Asp
 165 170 175

Phe Ile Ala Asp Asn Asn Asn His Tyr Lys Lys Phe Pro Gln Met Ser
 180 185 190

Ser Tyr Gln Arg Met Leu Val His Arg Val Ala Ala Tyr Phe Gly Leu
 195 200 205

16U 100 PCT.ST25

Asp His Asn Val Asp Gln Thr Gly Lys Ser Val Ile Ile Asn Lys Thr
 210 215 220
 Ser Ser Thr Arg Ile Pro Glu Gln Arg Phe Cys Glu His Leu Lys Asp
 225 230 235 240
 Glu Lys Gly Glu Glu Ser Gln Lys Arg Phe Ile Leu Lys Arg Asp Asn
 245 250 255
 Ser Ser Ile Asp Lys Glu Asp Asn Gln Ser Val Cys Ser Gln Glu Ser
 260 265 270
 Leu Phe Val Glu Asn Ser Arg Leu Leu Glu Asp Ser Asn Ile Cys Asn
 275 280 285
 Glu Thr Tyr Lys Lys Arg Gln Leu Phe Arg Gly Asn Arg Asp Gly Ser
 290 295 300
 Gly Arg Thr Ser Gly Ser Arg Gln Ser Ser Ser Glu Asn Glu Leu Lys
 305 310 315 320
 Trp Ser Asp His Gln Arg Ala Trp Ser Ser Thr Asp Ser Asp Ser Ser
 325 330 335
 Asn Arg Asn Leu Lys Pro Ala Met Thr Lys Thr Ala Ser Phe Gly Gly
 340 345 350
 Ile Thr Val Leu Thr Arg Gly Asp Ser Thr Ser Ser Thr Arg Ser Thr
 355 360 365
 Gly Lys Leu Ser Lys Ala Gly Ser Glu Ser Ser Ser Ser Ala Gly Ser
 370 375 380
 Ser Gly Ser Leu Ser Arg Thr His Pro Pro Leu Gln Ser Thr Pro Leu
 385 390 395 400
 Val Ser Gly Val Ala Ala Gly Ser Pro Gly Cys Val Pro Tyr Pro Glu
 405 410 415
 Asn Gly Ile Gly Gly Gln Val Ala Pro Ser Ser Thr Ser Tyr Ile Leu
 420 425 430
 Leu Pro Leu Glu Ala Ala Thr Gly Ile Pro Pro Gly Ser Ile Leu Leu
 435 440 445
 Asn Pro His Thr Gly Gln Pro Phe Val Asn Pro Asp Gly Thr Pro Ala
 450 455 460
 Ile Tyr Asn Pro Pro Thr Ser Gln Gln Pro Leu Arg Ser Ala Met Val
 465 470 475 480
 Gly Gln Ser Gln Gln Gln Pro Pro Gln Gln Gln Pro Ser Pro Gln Pro
 485 490 495
 Gln Gln Gln Val Gln Pro Pro Gln Pro Gln Met Ala Gly Pro Leu Val
 500 505 510
 Thr Gln Ser Val Gln Gly Leu Gln Ala Ser Ser Gln Ser Val Gln Tyr
 515 520 525

160 100 PCT.ST25

Pro Ala Val Ser Phe Pro Pro Gln His Leu Leu Pro Val Ser Pro Thr
 530 535 540

Gln His Phe Pro Met Arg Asp Asp Val Ala Thr Gln Phe Gly Gln Met
 545 550 555 560

Thr Leu Ser Arg Gln Ser Ser Gly Glu Thr Pro Glu Pro Pro Ser Gly
 565 570 575

Pro Val Tyr Pro Ser Ser Leu Met Pro Gln Pro Ala Gln Gln Pro Ser
 580 585 590

Tyr Val Ile Ala Ser Thr Gly Gln Gln Leu Pro Thr Gly Gly Phe Ser
 595 600 605

Gly Ser Gly Pro Pro Ile Ser Gln Gln Val Leu Gln Pro Pro Pro Ser
 610 615 620

Pro Gln Gly Phe Val Gln Gln Pro Pro Pro Ala Gln Met Pro Val Tyr
 625 630 635 640

Tyr Tyr Pro Ser Gly Gln Tyr Pro Thr Ser Thr Thr Gln Gln Tyr Arg
 645 650 655

Pro Met Ala Pro Val Gln Tyr Asn Ala Gln Arg Ser Gln Gln Met Pro
 660 665 670

Gln Ala Ala Gln Gln Ala Gly Tyr Gln Pro Val Leu Ser Gly Gln Gln
 675 680 685

Gly Phe Gln Gly Leu Ile Gly Val Gln Gln Pro Pro Gln Ser Gln Asn
 690 695 700

Val Ile Asn Asn Gln Gln Gly Thr Pro Val Gln Ser Val Met Val Ser
 705 710 715 720

Tyr Pro Thr Met Ser Ser Tyr Gln Val Pro Met Thr Gln Gly Ser Gln
 725 730 735

Gly Leu Pro Gln Gln Ser Tyr Gln Gln Pro Ile Met Leu Pro Asn Gln
 740 745 750

Ala Gly Gln Gly Ser Leu Pro Ala Thr Gly Met Pro Val Tyr Cys Asn
 755 760 765

Val Thr Pro Pro Thr Pro Gln Asn Asn Leu Arg Leu Ile Gly Pro His
 770 775 780

Cys Pro Ser Ser Thr Val Pro Val Met Ser Ala Ser Cys Arg Thr Asn
 785 790 795 800

Cys Ala Ser Met Ser Asn Ala Gly Trp Gln Val Lys Phe
 805 810

<210> 49
 <211> 3272
 <212> DNA
 <213> Homo sapiens

16U 100 PCT.ST25

<220>

<221> CDS

<222> (329)..(2710)

<223>

<400> 49

```

gtctattttt aatgctatatt aatgaaggag cgagcgcctc actcagcaat aaaagaagca      60
tgaggggaaga cagagcagtg catgggttatg gatactggac aaggatattt ggaaagggtg      120
acgatgtgtc acactgtgtg aggggaatcgc atggagatgg gcattccgaa ctgttaatgg      180
ggacatggga ctccagttgt ctctgatcac ttgtgtggat tttcctggcg tagaacgaca      240
gaagccgcta gtaagtcgcc aagacctaca gcaggaattc tgcaccaaag ggcataaaat      300
cttgttattt taatttgcatt ctggggaga atg tct gag caa gga gac ctg aat      352
                Met Ser Glu Gln Gly Asp Leu Asn
                1                    5

cag gca ata gca gag gaa gga ggg act gag cag gag acg gcc act cca      400
Gln Ala Ile Ala Glu Glu Gly Gly Thr Glu Gln Glu Thr Ala Thr Pro
    10                    15                    20

gag aac ggc att gtt aaa tca gaa agt ctg gat gaa gag gag aaa ctg      448
Glu Asn Gly Ile Val Lys Ser Glu Ser Leu Asp Glu Glu Glu Lys Leu
    25                    30                    35                    40

gaa ctg cag agg cgg ctg gag gct cag aat caa gaa aga aga aaa tcc      496
Glu Leu Gln Arg Arg Leu Glu Ala Gln Asn Gln Glu Arg Arg Lys Ser
    45                    50                    55

aag tca gga gca gga aaa ggt aaa ctg act cgc agt ctt gct gtc tgt      544
Lys Ser Gly Ala Gly Lys Gly Lys Leu Thr Arg Ser Leu Ala Val Cys
    60                    65                    70

gag gaa tct tct gcc aga cca gga ggt gaa agt ctt cag gat cag gaa      592
Glu Glu Ser Ser Ala Arg Pro Gly Gly Glu Ser Leu Gln Asp Gln Glu
    75                    80                    85

tca att cat tta cag ctt tcc agt ttt tcc agc ctg caa gag gag gat      640
Ser Ile His Leu Gln Leu Ser Ser Phe Ser Ser Leu Gln Glu Glu Asp
    90                    95                    100

aaa tct agg aaa gat gac tct gaa aga gaa aaa gaa aag gat aaa aac      688
Lys Ser Arg Lys Asp Asp Ser Glu Arg Glu Lys Glu Lys Asp Lys Asn
   105                    110                    115                    120

aaa gat aaa acc tct gaa aaa ccc aag atc aga atg tta tca aaa gat      736
Lys Asp Lys Thr Ser Glu Lys Pro Lys Ile Arg Met Leu Ser Lys Asp
   125                    130                    135

tgc agc caa gaa tac acg gat tct aca ggc ata gac tta cac gag ttt      784
Cys Ser Gln Glu Tyr Thr Asp Ser Thr Gly Ile Asp Leu His Glu Phe
   140                    145                    150

ctg att aac aca tta aag aat aat tcc agg gac agg atg ata ctt ttg      832
Leu Ile Asn Thr Leu Lys Asn Asn Ser Arg Asp Arg Met Ile Leu Leu
   155                    160                    165

aaa atg gag cag gaa att att gat ttc att gct gac aac aat aat cat      880
Lys Met Glu Gln Glu Ile Ile Asp Phe Ile Ala Asp Asn Asn Asn His
   170                    175                    180

tat aaa aag ttc cct cag atg tca tgc tat cag agg atg ctt gtc cat      928
Tyr Lys Lys Phe Pro Gln Met Ser Ser Tyr Gln Arg Met Leu Val His
   185                    190                    195                    200

cga gtg gca gct tat ttt gga ttg gat cac aat gtg gat caa aca gga      976
Arg Val Ala Ala Tyr Phe Gly Leu Asp His Asn Val Asp Gln Thr Gly
   205                    210                    215

aaa tct gtt atc atc aac aag acc agc agc acc aga ata cca gag caa      1024
Lys Ser Val Ile Ile Asn Lys Thr Ser Ser Thr Arg Ile Pro Glu Gln
   220                    225                    230

agg ttt tgt gaa cat tta aaa gat gaa aaa ggt gaa gaa tcc cag aag      1072
Arg Phe Cys Glu His Leu Lys Asp Glu Lys Gly Glu Glu Ser Gln Lys
   235                    240                    245

```

16U 100 PCT.ST25

cgg ttt atc ttg aag cga gat aac tct agt att gat aaa gaa gac aat Arg Phe Ile Leu Lys Arg Asp Asn Ser Ser Ile Asp Lys Glu Asp Asn 250 255 260	1120
cag tca gtt tgc tcc cag gaa agc ctt ttt gtg gaa aac agg ggc aac Gln Ser Val Cys Ser Gln Glu Ser Leu Phe Val Glu Asn Arg Gly Asn 265 270 275 280	1168
aga gat ggc tca ggg aga aca tct ggg agt cga cag agc agc tca gaa Arg Asp Gly Ser Gly Arg Thr Ser Gly Ser Arg Gln Ser Ser Ser Glu 285 290 295	1216
aat gaa ctc aag tgg tct gac cac caa agg gcc tgg agc agc aca gac Asn Glu Leu Lys Trp Ser Asp His Gln Arg Ala Trp Ser Ser Thr Asp 300 305 310	1264
tcc gac agt tcc aac cgc aat cta aag ccc gcc atg acc aag acg gcg Ser Asp Ser Ser Asn Arg Asn Leu Lys Pro Ala Met Thr Lys Thr Ala 315 320 325	1312
agt ttt ggg ggc atc acg gtg ctg acc agg ggt gac agc act tcc agt Ser Phe Gly Gly Ile Thr Val Leu Thr Arg Gly Asp Ser Thr Ser Ser 330 335 340	1360
act agg agt acc ggg aag ctg tcc aaa gca ggt tcc gag tct tcc agc Thr Arg Ser Thr Gly Lys Leu Ser Lys Ala Gly Ser Glu Ser Ser Ser 345 350 355 360	1408
agt gca ggc tcc tca gga tcg ctg tcc cgc acc cat cca cct ctc cag Ser Ala Gly Ser Ser Gly Ser Leu Ser Arg Thr His Pro Pro Leu Gln 365 370 375	1456
agc aca ccc cta gtc tca ggt gtg gca gct ggc tct cca ggc tgt gtg Ser Thr Pro Leu Val Ser Gly Val Ala Ala Gly Ser Pro Gly Cys Val 380 385 390	1504
cct tat cca gag aat gga ata ggg ggc cag gtt gct ccc agc agc acc Pro Tyr Pro Glu Asn Gly Ile Gly Gly Gln Val Ala Pro Ser Ser Thr 395 400 405	1552
agc tac atc ctc ctt cca ctt gaa gct gca aca ggc atc ccg cct gga Ser Tyr Ile Leu Leu Pro Leu Glu Ala Ala Thr Gly Ile Pro Pro Gly 410 415 420	1600
agc atc ctt ctt aat cca cac aca ggc cag ccc ttt gtg aat ccc gat Ser Ile Leu Leu Asn Pro His Thr Gly Gln Pro Phe Val Asn Pro Asp 425 430 435 440	1648
gga act cct gca ata tac aac cca ccc acc agt cag cag ccc ctg cga Gly Thr Pro Ala Ile Tyr Asn Pro Pro Thr Ser Gln Gln Pro Leu Arg 445 450 455	1696
agc gcc atg gtg ggg cag tcc caa cag cag ccg cca cag cag cag ccc Ser Ala Met Val Gly Gln Ser Gln Gln Pro Pro Gln Gln Gln Pro 460 465 470	1744
tcc ccg cag ccc caa cag cag gtc cag cca ccg cag cca cag atg gca Ser Pro Gln Pro Gln Gln Gln Val Gln Pro Pro Gln Pro Gln Met Ala 475 480 485	1792
ggc cct ctg gtc act cag tct gtc cag ggg ctg cag gct tcc tcc cag Gly Pro Leu Val Thr Gln Ser Val Gln Gly Leu Ala Ser Ser Gln 490 495 500	1840
tca gtg caa tat ccg gca gtc tct ttt cct ccc cag cac ctc cta cct Ser Val Gln Tyr Pro Ala Val Ser Phe Pro Pro Gln His Leu Leu Pro 505 510 515 520	1888
gtg tct cca acg cag cac ttt ccc atg aga gat gat gtg gca aca cag Val Ser Pro Thr Gln His Phe Pro Met Arg Asp Asp Val Ala Thr Gln 525 530 535	1936
ttt ggc cag atg acc ctg agc cgg cag tcc tcg ggg gag act cct gaa Phe Gly Gln Met Thr Leu Ser Arg Gln Ser Ser Gly Glu Thr Pro Glu 540 545 550	1984
ccc cca tca ggt cct gtc tac cca tcc tcc ctt atg cca cag ccg gcc Pro Pro Ser Gly Pro Val Tyr Pro Ser Ser Leu Met Pro Gln Pro Ala 555 560 565	2032

555	560	160 100 PCT.ST25 565	
cag cag ccc agc tat gta atc gcc tct aca ggc cag cag ctt cct aca Gln Gln Pro Ser Tyr Val Ile Ala Ser Thr Gly Gln Gln Leu Pro Thr 570 575 580			2080
gga gga ttc tca ggc tct ggc cct ccc atc tcc cag cag gtc ctc cag Gly Gly Phe Ser Gly Ser Gly Pro Pro Ile Ser Gln Gln Val Leu Gln 585 590 595 600			2128
ccc cct ccc tca cca cag gga tty gtg caa cag cct ccg cct gca cag Pro Pro Pro Ser Pro Gln Gly Phe Val Gln Gln Pro Pro Pro Ala Gln 605 610 615			2176
atg cct gta tat tat tac cca tct ggt cag tac cct acc tca acc acg Met Pro Val Tyr Tyr Tyr Pro Ser Gly Gln Tyr Pro Thr Ser Thr Thr 620 625 630			2224
caa cag tac cgg ccc atg gcc ccg gtt cag tac aac gct cag agg agt Gln Gln Tyr Arg Pro Met Ala Pro Val Gln Tyr Asn Ala Gln Arg Ser 635 640 645			2272
caa cag atg cca cag gca gca cag caa gca ggt tac cag cca gtc ttg Gln Gln Met Pro Gln Ala Ala Gln Gln Ala Gly Tyr Gln Pro Val Leu 650 655 660			2320
tct ggt caa cag gga ttc caa ggc cta ata gga gtg cag cag cca cct Ser Gly Gln Gln Gly Phe Gln Gly Leu Ile Gly Val Gln Gln Pro Pro 665 670 675 680			2368
cag agt cag aac gtg ata aat aac caa caa gga act ccg gtg caa agc Gln Ser Gln Asn Val Ile Asn Asn Gln Gln Gly Thr Pro Val Gln Ser 685 690 695			2416
gtg atg gtt tcc tac cca aca atg tct tct tat cag gtg cca atg acc Val Met Val Ser Tyr Pro Thr Met Ser Ser Tyr Gln Val Pro Met Thr 700 705 710			2464
cag ggt tct caa gga ctg ccc cag cag tca tac caa cag cca atc atg Gln Gly Ser Gln Gly Leu Pro Gln Gln Ser Tyr Gln Gln Pro Ile Met 715 720 725			2512
cta cct aac cag gca ggt caa ggg tca ctc cca gcc act gga atg cct Leu Pro Asn Gln Ala Gly Gln Gly Ser Leu Pro Ala Thr Gly Met Pro 730 735 740			2560
gtt tac tgt aat gtc aca ccg ccc acc cct cag aac aac ctt agg ctg Val Tyr Cys Asn Val Thr Pro Pro Thr Pro Gln Asn Asn Leu Arg Leu 745 750 755 760			2608
att ggc cca cac tgc ccc tcc agc act gtc cca gtg atg tca gct agc Ile Gly Pro His Cys Pro Ser Ser Thr Val Pro Val Met Ser Ala Ser 765 770 775			2656
tgc aga aca aac tgt gca agt atg agc aat gct ggt tgg cag gtc aaa Cys Arg Thr Asn Cys Ala Ser Met Ser Asn Ala Gly Trp Gln Val Lys 780 785 790			2704
ttc tga gagctctggc tgtggtacat ttcttcagat atttctcatg gcctttgatg Phe			2760
gaagaggaac aaggtgggaa aactggctga ggacttaagt attcactcaa cactcaaag			2820
attgctgctg gtattctgta aaaartaaac aaagactaat atacacgtta gctggttaat			2880
ggtgcatatt tctgtcatgt ctgctaggta tgcctttata gcttagctag tgacatgaat			2940
tcatcaaggc aagattytct cctaccactg aataccactg ttagattat aatatcccta			3000
atttggatta gttttgtact ttgtgttgag tttgtgatgc taaaagtatt taaaaattat			3060
atactaaatc acattgtacc aaagctgtaa tggaaaagca aagaagaayt gatgaattga			3120
aggaataatt tatatacatc atagagtttt cttttttaat ggatatatac tgtattgtag			3180
tgtttaatca aaataaaact atttgacctt atggaggaag gtcattgttt taaaaaaaaa			3240
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa			3272

16U 100 PCT.ST25

<210> 50
 <211> 793
 <212> PRT
 <213> Homo sapiens

<400> 50

```

Met Ser Glu Gln Gly Asp Leu Asn Gln Ala Ile Ala Glu Glu Gly Gly
1          5          10          15

Thr Glu Gln Glu Thr Ala Thr Pro Glu Asn Gly Ile Val Lys Ser Glu
20          25          30

Ser Leu Asp Glu Glu Glu Lys Leu Glu Leu Gln Arg Arg Leu Glu Ala
35          40          45

Gln Asn Gln Glu Arg Arg Lys Ser Lys Ser Gly Ala Gly Lys Gly Lys
50          55          60

Leu Thr Arg Ser Leu Ala Val Cys Glu Glu Ser Ser Ala Arg Pro Gly
65          70          75          80

Gly Glu Ser Leu Gln Asp Gln Glu Ser Ile His Leu Gln Leu Ser Ser
85          90          95

Phe Ser Ser Leu Gln Glu Glu Asp Lys Ser Arg Lys Asp Asp Ser Glu
100         105         110

Arg Glu Lys Glu Lys Asp Lys Asn Lys Asp Lys Thr Ser Glu Lys Pro
115         120         125

Lys Ile Arg Met Leu Ser Lys Asp Cys Ser Gln Glu Tyr Thr Asp Ser
130         135         140

Thr Gly Ile Asp Leu His Glu Phe Leu Ile Asn Thr Leu Lys Asn Asn
145         150         155         160

Ser Arg Asp Arg Met Ile Leu Leu Lys Met Glu Gln Glu Ile Ile Asp
165         170         175

Phe Ile Ala Asp Asn Asn Asn His Tyr Lys Lys Phe Pro Gln Met Ser
180         185         190

Ser Tyr Gln Arg Met Leu Val His Arg Val Ala Ala Tyr Phe Gly Leu
195         200         205

Asp His Asn Val Asp Gln Thr Gly Lys Ser Val Ile Ile Asn Lys Thr
210         215         220

Ser Ser Thr Arg Ile Pro Glu Gln Arg Phe Cys Glu His Leu Lys Asp
225         230         235         240

Glu Lys Gly Glu Glu Ser Gln Lys Arg Phe Ile Leu Lys Arg Asp Asn
245         250         255

Ser Ser Ile Asp Lys Glu Asp Asn Gln Ser Val Cys Ser Gln Glu Ser
260         265         270

Leu Phe Val Glu Asn Arg Gly Asn Arg Asp Gly Ser Gly Arg Thr Ser
275         280         285

```

16U 100 PCT.ST25

Gly Ser Arg Gln Ser Ser Ser Glu Asn Glu Leu Lys Trp Ser Asp His
 290 295 300
 Gln Arg Ala Trp Ser Ser Thr Asp Ser Asp Ser Ser Asn Arg Asn Leu
 305 310 315 320
 Lys Pro Ala Met Thr Lys Thr Ala Ser Phe Gly Gly Ile Thr Val Leu
 325 330 335
 Thr Arg Gly Asp Ser Thr Ser Ser Thr Arg Ser Thr Gly Lys Leu Ser
 340 345 350
 Lys Ala Gly Ser Glu Ser Ser Ser Ser Ala Gly Ser Ser Gly Ser Leu
 355 360 365
 Ser Arg Thr His Pro Pro Leu Gln Ser Thr Pro Leu Val Ser Gly Val
 370 375 380
 Ala Ala Gly Ser Pro Gly Cys Val Pro Tyr Pro Glu Asn Gly Ile Gly
 385 390 395 400
 Gly Gln Val Ala Pro Ser Ser Thr Ser Tyr Ile Leu Leu Pro Leu Glu
 405 410 415
 Ala Ala Thr Gly Ile Pro Pro Gly Ser Ile Leu Leu Asn Pro His Thr
 420 425 430
 Gly Gln Pro Phe Val Asn Pro Asp Gly Thr Pro Ala Ile Tyr Asn Pro
 435 440 445
 Pro Thr Ser Gln Gln Pro Leu Arg Ser Ala Met Val Gly Gln Ser Gln
 450 455 460
 Gln Gln Pro Pro Gln Gln Gln Pro Ser Pro Gln Pro Gln Gln Gln Val
 465 470 475 480
 Gln Pro Pro Gln Pro Gln Met Ala Gly Pro Leu Val Thr Gln Ser Val
 485 490 495
 Gln Gly Leu Gln Ala Ser Ser Gln Ser Val Gln Tyr Pro Ala Val Ser
 500 505 510
 Phe Pro Pro Gln His Leu Leu Pro Val Ser Pro Thr Gln His Phe Pro
 515 520 525
 Met Arg Asp Asp Val Ala Thr Gln Phe Gly Gln Met Thr Leu Ser Arg
 530 535 540
 Gln Ser Ser Gly Glu Thr Pro Glu Pro Pro Ser Gly Pro Val Tyr Pro
 545 550 555 560
 Ser Ser Leu Met Pro Gln Pro Ala Gln Gln Pro Ser Tyr Val Ile Ala
 565 570 575
 Ser Thr Gly Gln Gln Leu Pro Thr Gly Gly Phe Ser Gly Ser Gly Pro
 580 585 590
 Pro Ile Ser Gln Gln Val Leu Gln Pro Pro Pro Ser Pro Gln Gly Phe

16U 100 PCT.ST25
605

595 600

Val Gln Gln Pro Pro Pro Ala Gln Met Pro Val Tyr Tyr Tyr Pro Ser
610 615 620

Gly Gln Tyr Pro Thr Ser Thr Thr Gln Gln Tyr Arg Pro Met Ala Pro
625 630 635 640

Val Gln Tyr Asn Ala Gln Arg Ser Gln Gln Met Pro Gln Ala Ala Gln
645 650 655

Gln Ala Gly Tyr Gln Pro Val Leu Ser Gly Gln Gln Gly Phe Gln Gly
660 665 670

Leu Ile Gly Val Gln Gln Pro Pro Gln Ser Gln Asn Val Ile Asn Asn
675 680 685

Gln Gln Gly Thr Pro Val Gln Ser Val Met Val Ser Tyr Pro Thr Met
690 695 700

Ser Ser Tyr Gln Val Pro Met Thr Gln Gly Ser Gln Gly Leu Pro Gln
705 710 715 720

Gln Ser Tyr Gln Gln Pro Ile Met Leu Pro Asn Gln Ala Gly Gln Gly
725 730 735

Ser Leu Pro Ala Thr Gly Met Pro Val Tyr Cys Asn Val Thr Pro Pro
740 745 750

Thr Pro Gln Asn Asn Leu Arg Leu Ile Gly Pro His Cys Pro Ser Ser
755 760 765

Thr Val Pro Val Met Ser Ala Ser Cys Arg Thr Asn Cys Ala Ser Met
770 775 780

Ser Asn Ala Gly Trp Gln Val Lys Phe
785 790

<210> 51
<211> 1006
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (280)..(549)
<223>

<400> 51
gggcagcttg agacaggtgg agctggatca agctgtgaac gtgatttgct ggaagctggt 60
cattagtgtt gacgatgtgt cacactgtgt aagggaatcg catggagatg ggcattccga 120
actgttaatg gggacatggg actccagttg tctctgatca cttgtgtgga ttttcctggc 180
gtagaacgac agaagccgct agtaagtcgc caagacctac agcaggaatt ctgcaccaaa 240
gggcataaaa tcttgttatt ttaatttgca tctgggaga atg tct gag caa gga 294
Met Ser Glu Gln Gly
1 5
gac ctg aat cag gca ata gca gag gaa gga ggg act gag cag gag acg 342
Asp Leu Asn Gln Ala Ile Ala Glu Glu Gly Gly Thr Glu Gln Glu Thr
10 15 20
gcc act cca gag aac ggc att gtt aaa tca gaa agt ctg gat gaa gag 390

16U 100 PCT.ST25

Ala Thr Pro Glu Asn Gly Ile Val Lys Ser Glu Ser Leu Asp Glu Glu
 25 30 35

gag aaa ctg gaa ctg cag agg cgg ctg gag gct cag aat caa gaa aga 438
 Glu Lys Leu Glu Leu Gln Arg Arg Leu Glu Ala Gln Asn Gln Glu Arg
 40 45 50

aga aaa tcc aag tca gga gca gga aaa ggt aaa ctg act cgc agt ctt 486
 Arg Lys Ser Ser Lys Ser Gly Ala Gly Lys Gly Lys Leu Thr Arg Ser Leu
 55 60 65

gct gtc tgt gag gaa tct tct gcc aga cca gga ggt gaa agt ctt cag 534
 Ala Val Cys Glu Glu Ser Ser Ala Arg Pro Gly Gly Glu Ser Leu Gln
 70 75 80 85

gat cag act ctc tga aaactgcaaa tggaaaggaa ttcaaaagaa tttagattaa 589
 Asp Gln Thr Leu

aagttaaata aaaagtaggc acagtagtgc tgaattttcc tcaaaggctc tcttttgata 649
 aggctgaacc aaatataatc ccaagtatcc tctctccttc cttgttgagg atgtcttacc 709
 tctcagctcc caaaatgcac ttgcctataa gaaacacaat tgctggttca tatgaaactt 769
 wagaaatagt gaataagggtg catttaactt tggagaaata cttttatgsc tttggtggag 829
 atttctcaat actgcaaaag ttgtccagaa atgaatctga gctgatgggtg actttaagtt 889
 aatattatta atatatcact gcatattttt acccttattt ttgctcctta cagcaagatt 949
 agtaggttat aaaaatttaa atttaaacaa aattatttca tgacaaaatg ggaaact 1006

<210> 52
 <211> 89
 <212> PRT
 <213> Homo sapiens

<400> 52

Met Ser Glu Gln Gly Asp Leu Asn Gln Ala Ile Ala Glu Glu Gly Gly
 1 5 10 15

Thr Glu Gln Glu Thr Ala Thr Pro Glu Asn Gly Ile Val Lys Ser Glu
 20 25 30

Ser Leu Asp Glu Glu Glu Lys Leu Glu Leu Gln Arg Arg Leu Glu Ala
 35 40 45

Gln Asn Gln Glu Arg Arg Lys Ser Lys Ser Gly Ala Gly Lys Gly Lys
 50 55 60

Leu Thr Arg Ser Leu Ala Val Cys Glu Glu Ser Ser Ala Arg Pro Gly
 65 70 75 80

Gly Glu Ser Leu Gln Asp Gln Thr Leu
 85

<210> 53
 <211> 807
 <212> PRT
 <213> Mus musculus

<400> 53

Met Ser Glu Gln Gly Gly Leu Thr Pro Thr Ile Leu Glu Glu Gly Gln
 1 5 10 15

Thr Glu Pro Glu Ser Ala Pro Glu Asn Gly Ile Leu Lys Ser Glu Ser
 20 25 30

16U 100 PCT.ST25

Leu Asp Glu Glu Glu Lys Leu Glu Leu Gln Arg Arg Leu Ala Ala Gln
 35 40 45
 Asn Gln Glu Arg Arg Lys Ser Lys Ser Gly Ala Gly Lys Gly Lys Leu
 50 55 60
 Thr Arg Ser Leu Ala Val Cys Glu Glu Ser Ser Ala Arg Ser Gly Gly
 65 70 75 80
 Glu Ser His Gln Asp Gln Glu Ser Ile His Leu Gln Leu Ser Ser Phe
 85 90 95
 Pro Ser Leu Gln Glu Glu Asp Lys Ser Arg Lys Asp Asp Ser Glu Arg
 100 105 110
 Glu Lys Glu Lys Asp Lys Asn Arg Glu Lys Leu Ser Glu Arg Pro Lys
 115 120 125
 Ile Arg Met Leu Ser Lys Asp Cys Ser Gln Glu Tyr Thr Asp Ser Thr
 130 135 140
 Gly Ile Asp Leu His Gly Phe Leu Ile Asn Thr Leu Lys Asn Asn Ser
 145 150 155 160
 Arg Asp Arg Met Ile Leu Leu Lys Met Glu Gln Glu Met Ile Asp Phe
 165 170 175
 Ile Ala Asp Ser Asn Asn His Tyr Lys Lys Phe Pro Gln Met Ser Ser
 180 185 190
 Tyr Gln Arg Met Leu Val His Arg Val Ala Ala Tyr Phe Gly Leu Asp
 195 200 205
 His Asn Val Asp Gln Thr Gly Lys Ser Val Ile Ile Asn Lys Thr Ser
 210 215 220
 Ser Thr Arg Ile Pro Glu Gln Arg Phe Cys Glu His Leu Lys Asp Glu
 225 230 235 240
 Lys Ser Glu Glu Ser Gln Lys Arg Phe Ile Leu Lys Arg Asp Asn Ser
 245 250 255
 Ser Ile Asp Lys Glu Asp Asn Gln Asn Arg Met His Pro Phe Arg Asp
 260 265 270
 Asp Arg Arg Ser Lys Ser Ile Glu Glu Arg Glu Glu Glu Tyr Gln Arg
 275 280 285
 Val Arg Glu Arg Ile Phe Ala His Asp Ser Val Cys Ser Gln Glu Ser
 290 295 300
 Leu Phe Leu Asp Asn Ser Arg Leu Gln Glu Asp Met His Ile Cys Asn
 305 310 315 320
 Glu Thr Tyr Lys Lys Arg Gln Leu Phe Arg Ala His Arg Asp Ser Ser
 325 330 335
 Gly Arg Thr Ser Gly Ser Arg Gln Ser Ser Ser Glu Thr Glu Leu Arg
 340 345 350

16U 100 PCT.ST25

Trp Pro Asp His Gln Arg Ala Trp Ser Ser Thr Asp Ser Asp Ser Ser
 355 360 365
 Asn Arg Asn Leu Lys Pro Thr Met Thr Lys Thr Ala Ser Phe Gly Gly
 370 375 380
 Ile Thr Val Leu Thr Arg Gly Asp Ser Thr Ser Ser Thr Arg Ser Ala
 385 390 395 400
 Gly Lys Leu Ser Lys Thr Gly Ser Glu Ser Ser Ser Ser Ala Gly Ser
 405 410 415
 Ser Gly Ser Leu Ser Arg Thr His Pro Gln Ser Thr Ala Leu Thr Ser
 420 425 430
 Ser Val Ala Ala Gly Ser Pro Gly Cys Met Ala Tyr Ser Glu Asn Gly
 435 440 445
 Met Gly Gly Gln Val Pro Pro Ser Ser Thr Ser Tyr Ile Leu Leu Pro
 450 455 460
 Leu Glu Ser Ala Thr Gly Ile Pro Pro Gly Ser Ile Leu Leu Asn Pro
 465 470 475 480
 His Thr Gly Gln Pro Phe Val Asn Pro Asp Gly Thr Pro Ala Ile Tyr
 485 490 495
 Asn Pro Pro Gly Ser Gln Gln Thr Leu Arg Gly Thr Val Gly Gly Gln
 500 505 510
 Pro Gln Gln Pro Pro Gln Gln Gln Pro Ser Pro Gln Pro Gln Gln Gln
 515 520 525
 Val Gln Ala Ser Gln Pro Gln Met Ala Gly Pro Leu Val Thr Gln Arg
 530 535 540
 Glu Glu Leu Ala Ala Gln Phe Ser Gln Leu Ser Met Ser Arg Gln Ser
 545 550 555 560
 Ser Gly Asp Thr Pro Glu Pro Pro Ser Gly Thr Val Tyr Pro Ala Ser
 565 570 575
 Leu Leu Pro Gln Thr Ala Gln Pro Gln Ser Tyr Val Ile Thr Ser Ala
 580 585 590
 Gly Gln Gln Leu Ser Thr Gly Gly Phe Ser Asp Ser Gly Pro Pro Ile
 595 600 605
 Ser Gln Gln Val Leu Gln Ala Pro Pro Ser Pro Gln Gly Phe Val Gln
 610 615 620
 Gln Pro Pro Pro Ala Gln Met Ser Val Tyr Tyr Tyr Pro Ser Gly Gln
 625 630 635 640
 Tyr Pro Thr Ser Thr Ser Gln Gln Tyr Arg Pro Leu Ala Ser Val Gln
 645 650 655
 Tyr Ser Ala Gln Arg Ser Gln Gln Ile Pro Gln Thr Thr Gln Gln Ala

660 665 160 100 PCT.ST25
 670
 Gly Tyr Gln Pro Val Leu Ser Gly Gln Gln Gly Phe Gln Gly Met Met
 675 680 685
 Gly Val Gln Gln Ser Ala His Ser Gln Gly Val Met Ser Ser Gln Gln
 690 695 700
 Gly Ala Pro Val His Gly Val Met Val Ser Tyr Pro Thr Met Ser Ser
 705 710 715 720
 Tyr Gln Val Pro Met Thr Gln Gly Ser Gln Ala Val Pro Gln Gln Thr
 725 730 735
 Tyr Gln Pro Pro Ile Met Leu Pro Ser Gln Ala Gly Gln Gly Ser Leu
 740 745 750
 Pro Ala Thr Gly Met Pro Val Tyr Cys Asn Val Thr Pro Pro Asn Pro
 755 760 765
 Gln Asn Asn Leu Arg Leu Met Gly Pro His Cys Pro Ser Ser Thr Val
 770 775 780
 Pro Val Met Ser Ala Ser Cys Arg Thr Asn Cys Gly Asn Val Ser Asn
 785 790 795 800
 Ala Gly Trp Gln Val Lys Phe
 805

<210> 54
 <211> 648
 <212> PRT
 <213> Homo sapien

<400> 54

Met Ile Leu Leu Lys Met Glu Gln Glu Ile Ile Asp Phe Ile Ala Asp
 1 5 10 15
 Asn Asn Asn His Tyr Lys Lys Phe Pro Gln Met Ser Ser Tyr Gln Arg
 20 25 30
 Met Leu Val His Arg Val Ala Ala Tyr Phe Gly Leu Asp His Asn Val
 35 40 45
 Asp Gln Thr Gly Lys Ser Val Ile Ile Asn Lys Thr Ser Ser Thr Arg
 50 55 60
 Ile Pro Glu Gln Arg Phe Cys Glu His Leu Lys Asp Glu Lys Gly Glu
 65 70 75 80
 Glu Ser Gln Lys Arg Phe Ile Leu Lys Arg Asp Asn Ser Ser Ile Asp
 85 90 95
 Lys Glu Asp Asn Gln Ser Val Cys Ser Gln Glu Ser Leu Phe Val Glu
 100 105 110
 Asn Arg Leu Leu Glu Asp Ser Asn Ile Cys Asn Glu Thr Tyr Lys Lys
 115 120 125
 Arg Gln Leu Phe Arg Gly Asn Arg Asp Gly Ser Gly Arg Thr Ser Gly

Page 86

Ile Ser Gln Gln Val Leu Gln Pro Pro Pro Ser Pro Gln Gly Phe Val
 450 455 460 160 100 PCT.ST25

Gln Gln Pro Pro Pro Ala Gln Met Pro Val Tyr Tyr Tyr Pro Ser Gly
 465 470 475 480

Gln Tyr Pro Thr Ser Thr Thr Gln Gln Tyr Arg Pro Met Ala Pro Val
 485 490 495

Gln Tyr Asn Ala Gln Arg Ser Gln Gln Met Pro Gln Ala Ala Gln Gln
 500 505 510

Ala Gly Tyr Gln Pro Val Leu Ser Gly Gln Gln Gly Phe Gln Gly Leu
 515 520 525

Ile Gly Val Gln Gln Pro Pro Gln Ser Gln Asn Val Ile Asn Asn Gln
 530 535 540

Gln Gly Thr Pro Val Gln Ser Val Met Val Ser Tyr Pro Thr Met Ser
 545 550 555 560

Ser Tyr Gln Val Pro Met Thr Gln Gly Ser Gln Gly Leu Pro Gln Gln
 565 570 575

Ser Tyr Gln Gln Pro Ile Met Leu Pro Asn Gln Ala Gly Gln Gly Ser
 580 585 590

Leu Pro Ala Thr Gly Met Pro Val Tyr Cys Asn Val Thr Pro Pro Thr
 595 600 605

Pro Gln Asn Asn Leu Arg Leu Ile Gly Pro His Cys Pro Ser Ser Thr
 610 615 620

Val Pro Val Met Ser Ala Ser Cys Arg Thr Asn Cys Ala Ser Met Ser
 625 630 635 640

Asn Ala Gly Trp Gln Val Lys Phe
 645

<210> 55
 <211> 651
 <212> PRT
 <213> Homo sapien

<400> 55

Arg Asp Arg Met Ile Leu Leu Lys Met Glu Gln Glu Ile Ile Asp Phe
 1 5 10 15

Ile Ala Asp Asn Asn Asn His Tyr Lys Lys Phe Pro Gln Met Ser Ser
 20 25 30

Tyr Gln Arg Met Leu Val His Arg Val Ala Ala Tyr Phe Gly Leu Asp
 35 40 45

His Asn Val Asp Gln Thr Gly Lys Ser Val Ile Ile Asn Lys Thr Ser
 50 55 60

Ser Thr Arg Ile Pro Glu Gln Arg Phe Cys Glu His Leu Lys Asp Glu
 65 70 75 80

160 100 PCT.ST25
 Lys Gly Glu Glu Ser Gln Lys Arg Phe Ile Leu Lys Arg Asp Asn Ser
 85 90 95
 Ser Ile Asp Lys Glu Asp Asn Gln Ser Val Cys Ser Gln Glu Ser Leu
 100 105 110
 Phe Val Glu Asn Arg Leu Leu Glu Asp Ser Asn Ile Cys Asn Glu Thr
 115 120 125
 Tyr Lys Lys Arg Gln Leu Phe Arg Gly Asn Arg Asp Gly Ser Gly Arg
 130 135 140
 Thr Ser Gly Ser Arg Gln Ser Ser Ser Glu Asn Glu Leu Lys Trp Ser
 145 150 155 160
 Asp His Gln Arg Ala Trp Ser Ser Thr Asp Ser Asp Ser Ser Asn Arg
 165 170 175
 Asn Leu Lys Pro Ala Met Thr Lys Thr Ala Ser Phe Gly Gly Ile Thr
 180 185 190
 Val Leu Thr Arg Gly Asp Ser Thr Ser Ser Thr Arg Ser Thr Gly Lys
 195 200 205
 Leu Ser Lys Ala Gly Ser Glu Ser Ser Ser Ser Ala Gly Ser Ser Gly
 210 215 220
 Ser Leu Ser Arg Thr His Pro Pro Leu Gln Ser Thr Pro Leu Val Ser
 225 230 235 240
 Gly Val Ala Ala Gly Ser Pro Gly Cys Val Pro Tyr Pro Glu Asn Gly
 245 250 255
 Ile Gly Gly Gln Val Ala Pro Ser Ser Thr Ser Tyr Ile Leu Leu Pro
 260 265 270
 Leu Glu Ala Ala Thr Gly Ile Pro Pro Gly Ser Ile Leu Leu Asn Pro
 275 280 285
 His Thr Gly Gln Pro Phe Val Asn Pro Asp Gly Thr Pro Ala Ile Tyr
 290 295 300
 Asn Pro Pro Thr Ser Gln Gln Pro Leu Arg Ser Ala Met Val Gly Gln
 305 310 315 320
 Ser Gln Gln Gln Pro Pro Gln Gln Gln Pro Ser Pro Gln Pro Gln Gln
 325 330 335
 Gln Val Gln Pro Pro Gln Pro Gln Met Ala Gly Pro Leu Val Thr Gln
 340 345 350
 Ser Val Gln Gly Leu Gln Ala Ser Ser Gln Ser Val Gln Tyr Pro Ala
 355 360 365
 Val Ser Phe Pro Pro Gln His Leu Leu Pro Val Ser Pro Thr Gln His
 370 375 380
 Phe Pro Met Arg Asp Asp Val Ala Thr Gln Phe Gly Gln Met Thr Leu
 385 390 395 400

16U 100 PCT.ST25

Ser Arg Gln Ser Ser Gly Glu Thr Pro Glu Pro Pro Ser Gly Pro Val
 405 410 415
 Tyr Pro Ser Ser Leu Met Pro Gln Pro Ala Gln Gln Pro Ser Tyr Val
 420 425 430
 Ile Ala Ser Thr Gly Gln Gln Leu Pro Thr Gly Gly Phe Ser Gly Ser
 435 440 445
 Gly Pro Pro Ile Ser Gln Gln Val Leu Gln Pro Pro Pro Ser Pro Gln
 450 455 460
 Gly Phe Val Gln Gln Pro Pro Pro Ala Gln Met Pro Val Tyr Tyr Tyr
 465 470 475 480
 Pro Ser Gly Gln Tyr Pro Thr Ser Thr Thr Gln Gln Tyr Arg Pro Met
 485 490 495
 Ala Pro Val Gln Tyr Asn Ala Gln Arg Ser Gln Gln Met Pro Gln Ala
 500 505 510
 Ala Gln Gln Ala Gly Tyr Gln Pro Val Leu Ser Gly Gln Gln Gly Phe
 515 520 525
 Gln Gly Leu Ile Gly Val Gln Gln Pro Pro Gln Ser Gln Asn Val Ile
 530 535 540
 Asn Asn Gln Gln Gly Thr Pro Val Gln Ser Val Met Val Ser Tyr Pro
 545 550 555 560
 Thr Met Ser Ser Tyr Gln Val Pro Met Thr Gln Gly Ser Gln Gly Leu
 565 570 575
 Pro Gln Gln Ser Tyr Gln Gln Pro Ile Met Leu Pro Asn Gln Ala Gly
 580 585 590
 Gln Gly Ser Leu Pro Ala Thr Gly Met Pro Val Tyr Cys Asn Val Thr
 595 600 605
 Pro Pro Thr Pro Gln Asn Asn Leu Arg Leu Ile Gly Pro His Cys Pro
 610 615 620
 Ser Ser Thr Val Pro Val Met Ser Ala Ser Cys Arg Thr Asn Cys Ala
 625 630 635 640
 Ser Met Ser Asn Ala Gly Trp Gln Val Lys Phe
 645 650

<210> 56
 <211> 89
 <212> PRT
 <213> Homo sapien

<400> 56

Met Ser Glu Gln Gly Asp Leu Asn Gln Ala Ile Ala Glu Glu Gly Gly
 1 5 10 15

Thr Glu Gln Glu Thr Ala Thr Pro Glu Asn Gly Ile Val Lys Ser Glu
 20 25 30

16U 100 PCT.ST25

Ser Leu Asp Glu Glu Glu Lys Leu Glu Leu Gln Arg Arg Leu Glu Ala
 35 40 45

Gln Asn Gln Glu Arg Arg Lys Ser Lys Ser Gly Ala Gly Lys Gly Lys
 50 55 60

Leu Thr Arg Ser Leu Ala Val Cys Glu Glu Ser Ser Ala Arg Pro Gly
 65 70 75 80

Gly Glu Ser Leu Gln Asp Gln Thr Leu
 85

<210> 57
 <211> 88
 <212> PRT
 <213> Mus musculus

<400> 57

Met Ser Glu Gln Gly Gly Leu Thr Pro Thr Ile Leu Glu Glu Gly Gln
 1 5 10 15

Thr Glu Pro Glu Ser Ala Pro Glu Asn Gly Ile Leu Lys Ser Glu Ser
 20 25 30

Leu Asp Glu Glu Glu Lys Leu Glu Leu Gln Arg Arg Leu Ala Ala Gln
 35 40 45

Asn Gln Glu Arg Arg Lys Ser Lys Ser Gly Ala Gly Lys Gly Lys Leu
 50 55 60

Thr Arg Ser Leu Ala Val Cys Glu Glu Ser Ser Ala Arg Ser Gly Gly
 65 70 75 80

Glu Ser His Gln Asp Gln Thr Leu
 85

<210> 58
 <211> 4462
 <212> DNA
 <213> homo sapien

<220>
 <221> CDS
 <222> (1336)..(2163)
 <223>

<400> 58
 tctctttaat ctttgcctat ccaactggta tctcagtggt ctcttaaggc aaattccttt 60
 aatttttagtg acacactcat ttctactcga aacatttgcc tcattttcca tgatggacta 120
 ttttctaata caccacaat atttcctaga ttcttagtgg cagttcccta agtgtgccaa 180
 aatccccag cctgttagtg ttacttgta ttaagcaa atgattcaaa tcatcatttg 240
 atacatgatg gaaatccag gctcactcac caaatttcta gcaaatttg tgtgtgcttg 300
 taaggggat gggaggcag gagaaggcct tgtatttctt ctagtcatca caagctagt 360
 gtttttcttc ctctgtctgg aacttactcc ttgcacatac ctttttctgc actgtgccat 420
 catccacttc tctttacttc ctaactacca ccaactgaaa attatacata taaaagcttt 480
 aaacaaagtc tcttgaggct ctcaaggag tttacattac agtatagttc agcaacaat 540
 tttaaatcaa atagtacacc tctttattct tagaattccc tctgccaaaa aagaatcag 600

16U 100 PCT.ST25

```

ctactttttt ttaaattcaa ggtccaactt tctgttgtgt tgctgattct cttctctctt 660
tttccaacga cttccacttc tctctctagt ttacatgtct ccaaacctta agcctctgtt 720
aatactttca caataagtca attttgccaa cggtttgcct cccctagacc atctaggctg 780
ggcccagaac acctcatctt cactcccact gaagtgttcc tgaaggtcag ctctcactga 840
ccttgattct gctcccctac actgtcacca gaagctatcc acctatggtt ctaattcagt 900
aagtccaact ctctcacccc ctttttttgt ctcagctgtg tgggctttcc caggatggca 960
tgcaatggga cccctgtgcc atgcatattg taaaggaaaa tgccctccctc catgcgctac 1020
aaaacagcac atttatgatg gcactttgaa aagatatggg ttgtggtgtc acatattgac 1080
aattccttgg ccagaggctt aacagtgcc aagtgccag aagattaaga agacagcaaa 1140
aacagaaaag ggagaagatg gtgaagtagt tatataacat gagcgagaat gctcctgatt 1200
acaaagcaga gaaattgact ttttttctta gtgttttcta tagtcattgc tctatccctg 1260
ttctagaatt caagtcatga taagaatttc ttcacgttga cttcctgcat tgctttcaga 1320
cattgcaatt aaaga atg cga aga aag aac ctc aca gag gta aca gag ttt 1371
          Met Arg Arg Lys Asn Leu Thr Glu Val Thr Glu Phe
          1             5             10

gtt ttc ctg gga ttc tcc aga ttc cac aaa cat cac atc act ctc ttt 1419
Val Phe Leu Gly Phe Ser Arg Phe His Lys His His Ile Thr Leu Phe
          15             20             25

gtg gtt ttt ctc atc ctg tac aca tta act gtg gct ggc aat gcc atc 1467
Val Val Phe Leu Ile Leu Tyr Thr Leu Thr Val Ala Gly Asn Ala Ile
          30             35             40

atc atg acc atc atc tgc att gac cgt cac ctc cac act ccc atg tac 1515
Ile Met Thr Ile Ile Cys Ile Asp Arg His Leu His Thr Pro Met Tyr
          45             50             55             60

ttc ttc ctg agc atg ctg gct agc tca aag aca gtg tac aca ctg ttc 1563
Phe Phe Leu Ser Met Leu Ala Ser Ser Lys Thr Val Tyr Thr Leu Phe
          65             70             75

atc att cca cag atg ctc tcc agc ttc gta acc cag acc cag cca atc 1611
Ile Ile Pro Gln Met Leu Ser Ser Phe Val Thr Gln Thr Gln Pro Ile
          80             85             90

tcc cta gca ggt tgt acc acc caa acg ttc ttc ttt gtt acc ttg gcc 1659
Ser Leu Ala Gly Cys Thr Thr Gln Thr Phe Phe Phe Val Thr Leu Ala
          95             100             105

atc aac aat tgc ttc ttg ctc aca gtg atg ggc tat gac cac tat atg 1707
Ile Asn Asn Cys Phe Leu Thr Val Met Gly Tyr Asp His Tyr Met
          110             115             120

gcc atc tgc aat ccc ttg aga tac agg gtc att acg agc aag aag gtg 1755
Ala Ile Cys Asn Pro Leu Arg Tyr Arg Val Ile Thr Ser Lys Lys Val
          125             130             135             140

tgt gtc cag ctg gtg tgt gga gcc ttt agc att ggc ctg gcc atg gca 1803
Cys Val Gln Leu Val Cys Gly Ala Phe Ser Ile Gly Leu Ala Met Ala
          145             150             155

gct gtc cag gta aca tcc ata ttt acc tta cct ttt tgt cac acg gtg 1851
Ala Val Gln Val Thr Ser Ile Phe Thr Leu Pro Phe Cys His Thr Val
          160             165             170

gtt ggt cat ttc ttc tgt gac atc ctc cct gtc atg aaa ctc tcc tgt 1899
Val Gly His Phe Phe Cys Asp Ile Leu Pro Val Met Lys Leu Ser Cys
          175             180             185

att aat acc act atc aat gag ata atc aat ttt gtt gtc agg tta ttt 1947
Ile Asn Thr Thr Ile Asn Glu Ile Ile Asn Phe Val Val Arg Leu Phe
          190             195             200

gtc atc ctg gtc ccc atg ggt ctg gtc ttc atc tcc tat gtc ctc atc 1995
Val Ile Leu Val Pro Met Gly Leu Val Phe Ile Ser Tyr Val Leu Ile
          205             210             215             220

```

16U 100 PCT.ST25

atc tcc act gtc ctc aag att gcc tca gct gag ggt tgg aag aag acc 2043
 Ile Ser Thr Val Leu Lys Ile Ala Ser Ala Glu Gly Trp Lys Lys Thr
 225 230 235

ttt gcc acc tgt gcc ttc cac ctc act gtg gtc att gtc cat tat ggc 2091
 Phe Ala Thr Cys Ala Phe His Leu Thr Val Val Ile Val His Tyr Gly
 240 245 250

tgt gct tcc att gcc tac ctc atg ccc aag tca gaa aac tct ata gaa 2139
 Cys Ala Ser Ile Ala Tyr Leu Met Pro Lys Ser Glu Asn Ser Ile Glu
 255 260 265

caa gac ctc ctt ctc tca gtg acc taaaccatca tcaactccct gctgaaccct 2193
 Gln Asp Leu Leu Leu Ser Val Thr
 270 275

gttgtttaca gcctaaagaa caaggaggctc aaggatgccc tatgcagggc catgggcaga 2253
 aacattttctt aatgcattat tcctctatat aaatatacat ttagtcatag aatgtgtgt 2313
 ccttacttac attaaacaac cttacgactc tgtcccatgc agtctatgct gcaatgggat 2373
 gtgcatgtct tgctttggta tatttactac aaaatcttag tctctgtttc catatatttc 2433
 aaagttttgt ccaggcattt tcaactaggg atgtgagagg tcaaggagaa tgggcatgat 2493
 ttttaggaaa gagcatccaa atttctagga tgaagaaagg gacttttaaa agtatattaa 2553
 atatgattat attgtgttta aaaaataaaa agcaatgtgt ctcatTTTTg taatgcaatc 2613
 tacagaaat aaaactacaa aatcatcgca gcaaggtaag gatagtcaat aatgatggat 2673
 tcccttgaaa gaaaatagta tcagaattgt cagggaagg tgagatgagc gtattaaatt 2733
 taaaaaaaa taggaaagtt ggaaacact ggcttagtcc ttgaaaattt agtctttatt 2793
 attcaattta tgctaaagcc tttgctttta tccagtgtag tcgtcagatg ctggccatgg 2853
 ccacagatca tacttaactc tcacctttct aatctaaatt ccctaattga attctttctg 2913
 gctgctggtt ctctccatgg gatcaacttc tctctaataca ttatgaagaa aaattgagtt 2973
 ggtaagaag atctgtgccc tgtagaata agaaccataa aagctttcct catttgcaca 3033
 taccatggca cctcctggta gcataaagaa acaaaagtag aacaaaacaa acagtccag 3093
 cagtcaggag tagttcagaa gtataattgt agaactactc aattcaccaa aaaaggctaa 3153
 caagaaaaaa aaatatTTTT ccttagtaga ctcttttgag aaaaatatcc ttttccag 3213
 ggctttctgg gaactTTTT ggtttcatgc ctatacatc ctctatgtg gttatcttgg 3273
 gtgatgggtg atggtaggga gggtttaaga cttttgttga gtgcttcctt ctactaaaat 3333
 atttttgtc aagaatactt ctctgctcat tccctaatagc catTTTTctt ttttcaaga 3393
 cagtcatttt tttcttttcc ctccactgag aaagaatcga taacatccta aggagctcat 3453
 ctcaggtaaa gaatatcttt acagatttct gaattctaga ttggggagac cattgttctt 3513
 tcaaggctctg accagttctt tctaattcct gtcttgtgtc ttgcaattt ctacatcata 3573
 gaaacaaggc ttcctacaga agctgttggg ggctcaaagg ttgggtccaa gactcttggg 3633
 catgctatga ggtctttctt gacagcactc tcagggtcat cccactacag atctaaacct 3693
 ctatgcccaa caagagttag gcagattccc tccagaatga gagtgtctct cctgacagta 3753
 tgaaagaacc cctccactg ccataaaggc cacctgtact gccaaactctt aattataacc 3813
 tgcccttctc aattctctct ggctagcaac tacttgtaca tcctcatacc tgtctttttg 3873
 tgtgtagaag ggaagagatg agacagagag ggctgttctg agcaggaaag tacaatgatg 3933
 gaagtgggt ggtgatcaaa cctgttctat cctgtccag tggcagtggt caactccac 3993
 tgcccattac tctctttct acttctgcac acagccactc tccttcatta gttcatgaat 4053

16U 100 PCT.ST25
 caaagacaaa ggtttctcca gtattgtctc tacatctaaa tgctgcaaca gcagacatac 4113
 cacacgccac tcgtgcaatc aaaagttaaa tgtcacagcg gggcccacaa agagctggca 4173
 gaaaatatga gtcataaatc ttgaagggtga aatatctttg tctctaactc tgttaccac 4233
 aatggttcaa tcatatgtcc agtttttttt aaataaacat gtgtgcttct aatttcttct 4293
 tgacaataag atgttgctctt acatgggtgga tagaatagag gtcccagagt cagatgggtg 4353
 ctgagttcac ttcctcactc agcaccatc agaacccttt tatgaatata atgttttta 4413
 ttctgcctcg cttgtcatca aaagcatttt ttaatctcct taccattt 4462

<210> 59
 <211> 276
 <212> PRT
 <213> homo sapien

<400> 59

Met Arg Arg Lys Asn Leu Thr Glu Val Thr Glu Phe Val Phe Leu Gly
 1 5 10 15

Phe Ser Arg Phe His Lys His His Ile Thr Leu Phe Val Val Phe Leu
 20 25 30

Ile Leu Tyr Thr Leu Thr Val Ala Gly Asn Ala Ile Ile Met Thr Ile
 35 40 45

Ile Cys Ile Asp Arg His Leu His Thr Pro Met Tyr Phe Phe Leu Ser
 50 55 60

Met Leu Ala Ser Ser Lys Thr Val Tyr Thr Leu Phe Ile Ile Pro Gln
 65 70 75 80

Met Leu Ser Ser Phe Val Thr Gln Thr Gln Pro Ile Ser Leu Ala Gly
 85 90 95

Cys Thr Thr Gln Thr Phe Phe Phe Val Thr Leu Ala Ile Asn Asn Cys
 100 105 110

Phe Leu Leu Thr Val Met Gly Tyr Asp His Tyr Met Ala Ile Cys Asn
 115 120 125

Pro Leu Arg Tyr Arg Val Ile Thr Ser Lys Lys Val Cys Val Gln Leu
 130 135 140

Val Cys Gly Ala Phe Ser Ile Gly Leu Ala Met Ala Ala Val Gln Val
 145 150 155 160

Thr Ser Ile Phe Thr Leu Pro Phe Cys His Thr Val Val Gly His Phe
 165 170 175

Phe Cys Asp Ile Leu Pro Val Met Lys Leu Ser Cys Ile Asn Thr Thr
 180 185 190

Ile Asn Glu Ile Ile Asn Phe Val Val Arg Leu Phe Val Ile Leu Val
 195 200 205

Pro Met Gly Leu Val Phe Ile Ser Tyr Val Leu Ile Ile Ser Thr Val
 210 215 220

Leu Lys Ile Ala Ser Ala Glu Gly Trp Lys Lys Thr Phe Ala Thr Cys

160 100 PCT.ST25
 225 230 235 240
 Ala Phe His Leu Thr Val Val Ile Val His Tyr Gly Cys Ala Ser Ile
 245 250 255
 Ala Tyr Leu Met Pro Lys Ser Glu Asn Ser Ile Glu Gln Asp Leu Leu
 260 265 270
 Leu Ser Val Thr
 275

